VOL. XVIII JUNE, 1934 No. 3

THE JOURNAL OF APPLIED PSYCHOLOGY

JAMES P. PORTER, *Editor*
OHIO UNIVERSITY ATHENS, OHIO
With a Board of Cooperating Editors

SOME INTERESTING CONTRIBUTIONS IN THIS ISSUE:

- YALE NATHANSON and EDWARD M. TWITMYER. The Value of Specific Purpose in Advertising
- T. M. CARTER. The Effect of College Fraternities on Scholarship
- LOUIS F. HILLMAN. Training for Vocations in Those of High and Low Intelligence
- ROBERTA CRUTCHER. An Experimental Study of Persistence
- ALVHH R. LAUER and HAROLD L. KOTVIS. Automotive Manipulation in Relation to Vision
- F. H. FINCH and C. L. NEMZEK. Prediction of College Achievement from Data Collected During the Secondary School Period

PUBLISHED BI-MONTHLY BY THE EDITOR
OHIO UNIVERSITY, ATHENS, OHIO
PRINTED BY R. R. DONNELLEY & SONS COMPANY, CRAWFORDSVILLE, INDIANA

Entered as second-class matter at the post office at Athens, Ohio, under the act of March 3, 1879. Additional entry as second-class matter at the post office, Crawfordsville, Indiana.
Copyright 1934 by James P. Porter

The Journal of Applied Psychology

CO-OPERATING EDITORS

W. F. Book, Indiana University, W. V. Bingham, Personnel Research Federation, New York; M. E. Broom, State Teachers College, San Diego, Calif.; Robert H. Gault, Northwestern University; William Healy, Baker Foundation, Juvenile Court, Boston; H. L. Hollingworth, Columbia University; J. E. W. Wallin, Atlantic University; Herbert S. Langfeld, Princeton University; James H. Leuba, Bryn Mawr College; Henry M. McComas, Johns Hopkins University; Walter D. Scott, Northwestern University; C. E. Seashore, University of Iowa; Daniel Starch, New York; Edward K. Strong, Jr., Leland Stanford Jr. University; L. M. Terman, Leland Stanford Jr. University; E. L. Thorndike, Columbia University; Joseph Jastrow, New School of Social Research, New York; G. M. Whipple, Danvers, Mass.; Helen Thompson Woolley, Teachers College, Columbia University.

The scope of The Journal of Applied Psychology will include the following: (a) The application of psychology to vocational activities, such as law, art, public speaking, industrial and commercial work, and problems of business appeals. (b) Studies of individual mentalities, such as types of character, special talents, genius, and individual differences, including the problems of mental diagnosis and vocational prognosis. (c) The influence of general environment conditions, such as climate, weather, humidity, temperature; also such conditions as nutrition, fatigue, etc. (d) The psychology of everyday activities, such as reading, writing, speaking, singing, playing games or playing musical instruments, sports, etc.

Communications, manuscripts, reviews, summaries, notices, books and offprints for review should be addressed to Dr. James P. Porter, Ohio University, Athens, Ohio. Contributors are requested to follow the conventional method of citation employed by this Journal; in every case the year, volume and pages should be specified.

The Journal of Applied Psychology is published bi-monthly. The subscription price is $5.00 a year. Single numbers $1.00 to $2.00, depending upon the number of pages. Remittances and business communications should be addressed to the Editor, James P. Porter, Ohio University, Athens, Ohio.

Vol. XVIII JUNE, 1934 No. 3

THE JOURNAL OF APPLIED PSYCHOLOGY

Edited by

James P. Porter, Ohio University
Athens, Ohio

TABLE OF CONTENTS

YALE NATHANSON AND EDWARD M. TWITMYER. The Value of Specific Purport in Advertising 319

CARNEY LANDIS AND S. E. KATZ. The Validity of Certain Questions which Purport to Measure Neurotic Tendencies 343

M. E. BROOM. A Study of Literary Appreciation 357

S. N. STEVENS AND E. F. WONDERLIC. The Relationship of the Number of Questions Missed on the Otis Mental Tests and the Ability to Handle Office Detail 364

PAUL A. WITTY AND VIOLA THEMAN. The Psycho-Educational Clinic 369

T. M. CARTER. The Effect of College Fraternities on Scholarship 393

LOUIS F. HILLMAN. Training for Vocations in Those of High and Low Intelligence 401

ROBERTA CRUTCHER. An Experimental Study of Persistence 409

LORRAINE B. ABEL. A Test of Eye-Hand Coordination at the Fifth Year Level 418

ALVHH R. LAUER AND HAROLD L. KOTVIS. Automotive Manipulation in Relation of Vision 422

MAY V. SEAGOE. An Evaluation of Certain Intelligence Tests. 433

HARRY TENWOLDE. A Comparison of the Handwriting of Pupils in Certain Elementary School Grades "Now and Yesterday" 438

VERNE W. LYON. A Study of Vocational Abilities of Students Who Have Attended the Illinois School for the Deaf... 443

F. H. FINCH AND C. L. NEMZEK. Prediction of College Achievement from Data Collected During the Secondary School Period 454

MARGUERITE R. HERTZ. The Reliability of the Rorschach Ink-Blot Test 461

JAMES PAGE. Introversion-Extroversion and the Functional Psychoses 478

EDNA WILLIS McELWEE. Recognition of Melodies and Tonal Memory 489

JOSEPH ZUBIN. Note on the Standard Error of the Difference between Coefficients of Variation of Correlated Variables 491

Notes and News 493

Book Reviews 496

New Books and Pamphlets 499

THE VALUE OF SPECIFIC PURPOSE IN ADVERTISING[1]

YALE NATHANSON and EDWARD M. TWITMYER

University of Pennsylvania

THE method of the questionnaire for acquiring information of individual reactions to a situation dates beyond Galton in the early 80's to John Locke's "Essay Concerning Human Understanding" 200 years before (1690). There are definite objections to the questionnaire method to be sure, but (1) proper selection of material, (2) effective type of interrogation, (3) correct interpretation, and (4) planned purpose, make possible the control of variables and the validation of results.

Individuals are inclined, unless the object of a test is somewhat obscured, to look for objectives and answer not what would be a purely natural, common sense reply, but to read meaning into material beyond what is originally intended, thus defeating the purpose of the analysis. Therefore, the method here employed made the test appear as one of a purely academic nature, a test in "Analytic Discrimination." In no instance was the goal idea, as such, stressed. This was averted by reading a brief series of instructions which was repeated in substance in the introduction to the test.

Several hundred more tests were distributed than were returned by the students. The students were urged to perform the test only if they felt so inclined and if their other activities did not press them so seriously as to interfere with their best judgments and their application to the problem at hand. The replies, upon which this analysis is based, are therefore the product of *interested persons,* whose primary object in the cooperation offered is that of scientific interest. Thus we are assured that in the instances where hypothetical questions were formulated and certain assumptions were of necessity accepted by the person tak-

[1] A psychological survey based on the method of the questionnaire.

ing the test, that these were consistently maintained and the test approached as an academic problem not unlike many other tests dealing with purely academic or factual content. The individual to whom the test was submitted and who finally exercised the time and effort to arrive at a solution represents an individual competent to understand the situations presented and exercise the amount of judgment necessary to make the conclusions valid.

SEX DIFFERENCES

The test results show that an almost equal number of males and females participated. In the case of the males the added note was furnished: "Assume that your judgment is based on the additional hypothesis that you are thinking in terms of purchase and selection for a female member of the family."

It is interesting to note that in the greater number of instances the male and female judgments coincide. Thus it shows clearly that there is no marked sex difference in the *fundamental appeals* which advertising makes regardless of the commodity presented. This is important when both males and females help make a decision. Were it otherwise, it would be necessary to present a series of advertisements in which the appeal would be alternately made to male and female groups. Under the present circumstances such a method is not indicated.

GEOGRAPHICAL DISTRIBUTION

The literature on tests shows certain sectional differences, dependent upon the various factors, racial characteristics, climate and other local environmental conditions. It is not uncommon, for instance, in making a survey of advertisements of national producers to observe that in different sections of the country the body of the advertisement is differently written to allow for these individual sectional differences. Even nomenclature and phraseology show these variations. Those participating in the present test are now living in Philadelphia and are subject to local influences, despite the fact that they are from many parts of the United States. Distribution according to locality follows:

Male

Philadelphia (including Suburbs)	130
Pennsylvania	31
New York	25
New Jersey	17
Connecticut	8
Ohio	4
Indiana	3
Delaware	2
Maryland	2
Missouri	2
Colorado	1
Kentucky	1
Louisiana	1
Nebraska	1
Canal Zone	1
Porto Rico	1
Egypt	1

Female

Philadelphia (including Suburbs)	235
Pennsylvania	57
New Jersey	38
New York	3
Maryland	5
Delaware	4
Massachusetts	3
West Virginia	2
Ohio	1
Connecticut	1

TIME

The problem of the amount of time normally spent in "attending" to an advertisement is highly important. Casual observations of persons on street cars and in other public places tend to support the belief that the initial arrest of attention, which invites a person to read further into an advertisement, is a very rapid process. Even the reading of the advertisement subsequently, unless there is some particular point of interest, does not

consume a great deal of time. To the average lay reader advertising is scanned rather hurriedly and therefore a superficial judgment of advertising by the lay reader cannot be accepted as possessing a great deal of validity. Therefore, in this survey we requested the participants to designate the approximate time which they spent on the advertisements. The male averages 21.4 minutes, and the female 22.2 minutes. There is less than a minute's difference shown by males and females, regardless of the fact that we did not indicate how long a period was necessary for an accurate analysis. This has several points of value from a psychological point of view. First, it means that the content and the presentation of the material were of such character that it required approximately twenty-one or twenty-two minutes, and secondly, we are convinced by checking the males against the females as to time of judgment that this amount of time was actually spent.

TIME

Male

1st sheet	161 cases	3405	minutes
2nd sheet	72 cases	1587	minutes
Total	233 cases	4992	minutes
Average time	21.4 minutes		

Female

1st sheet	159 cases	3475	minutes
2nd sheet	156 cases	3535	minutes
3rd sheet	31 cases	678	minutes
Total	346 cases	7688	minutes
Average time	22.2 minutes		

CONDITIONS OF TEST

The analyses were made by the students at home or in the laboratory only under such conditions as would be ideal for reflection and analysis. The instructions were very specific that all the tests should be completed at one sitting, that there was no need to hurry and that the time was not an element of the test. This was done for the reason that we were anxious to have an ontic uniformity as well as a group conformity. It is a difficult

thing, of course, to control the attitude of persons who submit to any kind of test, but we know from years of experience that one can establish a maximum amount of interest and accuracy on the part of participants by the proper kind of instruction and the proper stressing of the value and need of their scientific accuracy of judgment. We are confident that intellectual honesty was exercised and that the integrity and cooperation of the individuals likewise were demonstrated throughout.

AGE

The ages of the individuals run from that of the youngest students, eighteen or nineteen years of age, to many of the older students, who are teachers in the schools adjacent to Philadelphia, who come to the University for work on Saturdays and in the evenings, running into the forties. Thus, there is a very suitable spread which reaches practically all those persons of the age who are interested in the type of merchandise presented.

SOCIAL AND EDUCATIONAL STATUS

The persons who submitted the answers to the questionnaire belong to that group of individuals who play an important part in the selection of their own purchases. Every class of society is represented, from young men and women of very humble circumstances to those who are independently wealthy so that the distribution eliminates any possible skewing of a curve meant to show a distribution in terms of financial classes.

As regards educational status the very fact that these men and women are students at an educational institution is *prima facie* evidence that they have a certain known amount of academic material and training. The percentage which they represent in terms of the general populace educationally is rather small, so that if any criticism might be advanced, it could only be forwarded in the direction of the fact that these people are a selected few. However, to offset this there is to be considered the wide range of the homes which are represented. The importance of such a study as the present one is the psychological determination of those *basic factors* which actuate people in making their judgments. These many factors are dependent upon the competency

MASON & ADAMS

A Sale of Sample Furs At About Half Price

of interest to every woman who wants a Smart Fur Coat

She will be able to purchase it, in this Sale, at about half the original price! These coats, samples used on the road by our fashionable—are offered tomorrow at tremendous reductions. Therefore, this opportunity—coming just in time to meet your plans of "A Fur Coat for Christmas!"

◆◆◆

Pay $10 Deposit

A $10 deposit reserves Your Purchase. You can arrange a charge account and wear while paying.

Trade In Your Old Furs

Trade in your old Fur, our liberal allowance for them will reduce cost of your new Fur Coat.

Coats Under $100

Modest Coney	Brown Styles	$98.50
Leopard Coney	Smart Set+ Models	97.50
Brown Coney	Self-trimmed	95.00
Mink Marmot (Dyed	Barony, (Fitted), Shawl Models	97.50
Natural Raccoon	(Pouch) Swagger Models	97.50
French Seal	(Dyed Coney) Sweep Models	98.50
Silver Coney	Fur trimmed	98.50
Hair Seal	Australian Opossum trimmed	98.50
China Civet Cat	(Imperial Fur), Sweep Styles	98.50
Modern Beaver	(Dyed Coney, Youthful Models	97.50
French Seal	(Dyed Coney) Beaver Coney trim'd	97.50
French Seal	(Dyed Coney), Shoulder trimmed	98.50
Natural Opossum	Swagger Models	98.50
French Seal	(Dyed Coney), Flash trimmed	98.50
Russian Pony	Shawl trimmed	98.50
Wallaby Raccoon	Sports Models	98.50
Russian Pony	Fur trimmed	98.50
Platinum Carneal	Shawl Models	98.50
American Broadtail	Fur trimmed	98.50
Natural Pony	Reversible Model, Beach styling	98.50
Natural Muskrat	Shawl styles	97.50
Platinum Carneal	Fur trimmed	100.00

Coats Up to $200

Platinum Carneal	Flash trimmed	$110.00
Leopard Cat	Raccoon trimmed	110.00
Sea Lion	Civet Cat trimmed	110.00
Russian Pony	Squirrel trimmed	110.00
Kidskin	Shawl trimmed models	110.00
Natural Raccoon	Sports Models	110.00
Natural Muskrat	Reversible Models	110.00
Snowflake Weasel	Extra Squirrel trimmed	110.00
Antelope Caste	Shawl Styles	107.50
Dyed Gazelle	Fur trimmed	107.50
Natural Muskrat	Fur trimmed	105.00
Platinum Carneal	Civet Cat trimmed	105.00
White Coney	Frontier Models, Black Fox trim'd.	105.00
Modern Seal	(Dyed Muskrat), Shawl Models	107.50
Antelope Caste	Points trimmed	107.50
Black Bear	Coon, Squirrel trimmed	105.00
Modern Seal	(Dyed Muskrat), Leopard trimmed	105.00
Modern Seal	(Dyed Muskrat), Squirrel trimmed	105.00
Modern Seal	(Dyed Muskrat), Skunk trimmed	105.00
Civet Cat Caste	Sports Models	105.00
Platinum Carneal	Squirrel trimmed	105.00
Black Carneal	Squirrel trimmed	105.00
Natural Nutria	Shawl Models	105.00
Coon Seal	Fur trimmed	105.00
Platinum Red Carneal	Squirrel trimmed, Reversible Models	175.00
Leopard	Red Fox trimmed	105.00
Russian Leopard	Wolf Fox trimmed	105.00
Alaska Caste	Yoshi Squirrel trimmed	105.00
Jap Mink	Fox trimmed	105.00

Coats Up to $400

Modern Seal	(Dyed Muskrat), Reversible Mod's	$225.00
American Broadtail	Fox trimmed	225.00
Platinum Carneal	Platinum Fox trimmed	225.00
Natural Muskrat	Platinum Fox trimmed	225.00
Modern Seal	(Dyed Muskrat), Silver Fox trim'd	275.00
Gray Kolinsky	Shawl Models	297.50
Indian Barony Dyed	Reversible Models	295.00
Alaska Seal	Shawl Styles	225.00
Alaska Seal	Sports Models	225.00
Northern Barony Carneal	Fox trimmed	295.00
Alaska Seal	Mink trimmed	275.00
Gray Kolinsky	Silver Fox trimmed	295.00

Coats Over $400

White Ermine Cape	White Fox trimmed	$450.00
Cocoa Ermine	Jones Marten trimmed	545.00
Brown Mahler Carneal	Brown Marten trimmed	545.00
Natural Mink	Shawl Silver Fox Styles	595.00
Platinum Ermine	Platinum Fox trimmed	595.00
Cocoa Ermine	Cocoa Fox trimmed	595.00
Natural Mink	Finely Matched Silver	595.00
White Ermine Cape	White Mid'ng trimmed	595.00
Black Carneal	Brown Marten trimmed	595.00
Silver Ermine	Platinum Fox trimmed	595.00
Polar Ermine	Polar Fox trimmed	595.00
Natural Mink	Natural Mink trimmed	1195.00
Natural Mink	Finely Matched Silver	1195.00
Natural Mink	Colonial Paws	995.00

Coats Pictured

A. Platinum Ermine	Platinum Fox trimmed	$95.00
B. Mink Marmot	(Dyed Marmot), Shawl Models	97.50
C. Alaska Beaver	Sports Model	98.50
D. Russian Pony	Fur trimmed	98.50
E. Natural Raccoon	Swagger Models	98.50
F. French Seal	(Dyed Coney), Shawl Models	98.50
G. Jap Mink	Fox trimmed, Latest styles	98.50

MASON & ADAMS

A $10 deposit reserves your purchase; arrange a charge account and wear while paying.

Trade in your old furs. Our liberal allowance will reduce the cost of your new Fur Coat.

EXCLUSIVE FUR FASHIONS

For the Woman Whose Discrimination Is Evident at First Glance and Whose Good Taste Is Never Uncertain

Exquisite beauty, paramount quality, the distinct originality of design characterize these lustrous fur garments offered by America's largest retail Fur establishment. These values at $495 are made possible by our 37th Anniversary Celebration, which commemorates the achievements of four generations of business endeavor. The confidence which has been ours for so long will be perpetuated throughout this great Sale and the values we offer will emphasize our Fur leadership and create new high standards of Fur value.

Choose from the Following at

$495

African Baby Leopard Beaver trimmed	**Natural Mink** Shawl collar
Platinum Mahler Carneal Platinum Fox trimmed	**Jap Mink** Let-Out styles
White Ermine White Fox trimmed	**Gray Kolinsky** Flower-Back models

One Day Clearance of 178 Fine Fur Coats

MASON & ADAMS — SATURDAY

At Less Than Wholesale Cost !

Brown Moire Caracul Baum Marten, Trimmed $287

THESE ARE THE FACTS—We are offering these 178 coats at tremendous reductions in order to start our new year with a clean and complete stock. These models are the smartest coats of the season and have, therefore, been sold down to one or two of a kind. And you can secure any one now AT LESS THAN TODAY'S WHOLESALE COST OF PELTS ALONE!

Last week more than 800 women bought their Fur Coats for next season here. Why? Because they could get a $100 Fur Coat for $65—a $200 Fur Coat for $130—a $600 Coat for $450. Because they could buy on our Small Deposit Budget Plan; because they could trade in their old Furs; because their Furs will be stored free of charge; because they have the spring and summer to pay for them.

THIS IS YOUR OPPORTUNITY — do not pass it by. Choose your next season's Fur Coat here TOMORROW

12	Mink Marmot	Newest Models	*58 Coats That Were to $145*
8	French Seal	Mink Marmot Trimmed	
5	Antelope	Sports Models	
13	Russian Pony	Beaver and Fox Trimmed	
9	Natural Opossum	Sports Models	
11	Brown and Platinum Caracul	Fox Trimmed	**$74**
12	Natural Raccoon	Sports Models	*52 Coats That Were to $225*
10	Australian Opossum	Jaunty Styles	
4	Hudson Seal	Skunk Trimmed	
2	Hair Seal	Tailored Models	
14	Natural Muskrat	Opossum Models	
9	Pony Coats	Fitch Trimmed	**$125**
6	Jap Mink	Fox Trimmed	*30 Coats That Were to $330*
6	Cocoa Squirrel	Fox Trimmed	
4	Natural Krimmer	Jaunty Styles	
3	Natural Squirrel	New Models	
8	Hudson Seal	Squirrel Trimmed	
10	Black and Platinum Caracul	Fox Trimmed	**$175**
4	Alaska Beaver	Jaunty Styles	*29 Coats That Were to $425*
6	Natural Squirrel	Platinum Fox Trimmed	
4	African Leopard	Fitch Trimmed	
6	Platinum Moire Caracul	Platinum Fox Trimmed	
7	Hudson Seal	Stout's Trimmed Models	
2	Brown Moire Caracul	Baum Marten Trimmed	**$287**

Buy Now for Next Season

A $10 deposit reserves your purchase until next season, payments continuing weekly or monthly during the Spring and Summer.

If you desire to wear coat now, open a charging account with us and wear coat while paying.

Trade In Your Old Furs

Old Furs accepted in part payment for new coats. Wear the old Furs until the end of the season, while your purchase is reserved. Trade-In Manager will gladly give estimate.

You Save From $50 to $125 on Every Coat

French Seal is Dyed Coney—Hudson Seal is Dyed Muskrat—Mink Marmot is Dyed Marmot

EXTRA SPECIAL
5 Mink Coats
Specially Reduced from $1500 to
$895

of the individual's heredity, his environment, as reflected in those affective qualities which are the resultants of instinct, emotion and, finally, his native capacity for ratiocination.

The tests given each participant consisted of a large envelope containing four sheets, 8½" x 11", the material of which is presented in the discussion and three advertisements, "A," "B," "C."

The questionnaire, instructions, results and discussion follow:

ADVERTISING ANALYSIS T-N-25

A scientific attitude is essential to any valid analysis. In answering the questions of this test T-N-25, ADVERTISING ANALYSIS, the examiner is confident of your absolute cooperation and intellectual integrity in arriving at your judgments.

Do not invite the opinions of others but make this an individual determination. Thorough application, intelligent reading of the questions and real concentration will mean a limited amount of time for its completion.

The questions are not difficult, the interpretation not involved. The only caution is to exercise care and intellectual honesty. Do not skip any questions but answer them in the order asked, and keep a record of how long a time the entire test requires. Do not hurry. Time is not an element of the test. Complete test at one sitting.

(NOTE: In the case of male examinees assume that your judgments are based on the additional hypothesis that you are thinking in terms of purchase and selection for a female member of the family.)

Problem I reads as follows:

You have often observed advertisements concerning various merchandise (furs, dresses, hats, gloves, shoes, jewelry, etc.).

Place the figure one (1) after the factor below which you feel is most decisive in causing you to purchase the particular item.

Place the figure two (2) after what you consider the next factor of importance.

Place the figure three (3) after what you consider the least deciding factor.

DESCRIPTION OF ARTICLE.....

NAME OF FIRM.....

PRICE (AMOUNT OF MONEY INVOLVED)....

The primary object in Problem I was to determine which type of advertisement was most effective. It is sometimes difficult to weigh one specific factor against another because individuals are of such a complex character and nature that not one primary emotion or instinct or sensation becomes involved, but a series of factors superimposed one upon the other. Thus, in Problem I, an attempt was made to have the judgment established on the basis of a verbal connotation, a description, verbal suggestion not concrete illustration, from which a purer judgment was possible. It is difficult to appreciate what this purity of material means in the isolation of one process from another and one sensation from that suggested by another. It is not possible to break up all the competencies of the human being into a number of specific abilities, and then to test and evaluate each of these abilities regardless of the others without overlapping. The behavior output of any human being is a complex of abilities.

The three types of judgment referred to were:

(1) Description of article—"Merchandising Appeal"
(2) Name of Firm—"Institutional Appeal"
(3) Price—"Price Appeal"

Males and females both conclusively indicate that the description of the article (merchandising) is the most decisive factor in purchase. The next important factor is "Price Appeal." The "Institutional Advertisement," emphasizing name of firm ranks third in making a determination. This is highly significant in view of the fact that even with the highly selective group of individuals participating in the test, the institutional appeal is third in importance. The idea has been advanced that the institutional appeal has more weight, but a group of persons accustomed to the respect and recognition of authority, either in academic matters or in the general world of things, still revert to the more in-

stinctive, more habitual and economically necessary appeal made by the description of an article or by its price, and of these two, merchandising is the more impressive factor.

RESULTS

PROBLEM I

		Male		*Female*		*Total*	
	1.	99		161		260	
A	2.	53		70		123	
	3.	83	235	118	349	201	584
	1.	68		102		170	
B	2.	59		77		136	
	3.	108	235	170	349	278	584
	1.	68		86		154	
C	2.	123		202		325	
	3.	44	235	61	349	105	584

The following questions all require constant reference to the three advertisements contained herein and designated "A," "B," "C." Place them before you in full view—do not trust to memory.

PROBLEM II

Problem II reads as follows:

To answer Problem II, you must assume the following conditions:

You must have a *new coat*

You are not decided on the *kind of coat*

You have *not* determined the amount you will spend for the new coat

You have *not* determined where you are going to purchase the new coat

In view of these assumptions, answer the following question:

Which of the advertisements furnished you (A, B, C) in their rank order most intensifies your desire to purchase a *fur coat*

1st choice —— 2nd choice —— 3rd choice ——

In each space above provided for choice write A, B, or C. e.g. If the advertisement were designated X, Y, Z instead of A, B, C, your answer might appear:

1st choice __z__ 2nd choice __y__ 3rd choice __x__

In the next several problems of the questionnaire an attempt was made to maintain the same number of constants in each instance, introducing only one variable at a time. Thus, the assumption offered in Problem II is the necessity for a new coat, with a decision as to kind, amount of money involved, place of purchase not known. There is involved the factor of motivation, desire or urge. Hess very ably says that the function of the advertiser is to stimulate desire "on the part of the class or classes for which his proposition is intended and to consider himself a factor in the changing of an economic system, so that all who would be really benefitted should have the opportunity of rightly buying what he has to sell." Certainly the first situation is contained here, at least, in regard to the money and place of purchase. The original assumption in Problem II is the necessity for a new coat. Whether or not that new coat shall be of fur or cloth, or the amount of money to be spent, or the place where the purchase is to be made, is not the primary consideration of the advertisement as it applies to this particular problem. There is an altogether different psychological appeal where the problem is one of stimulating desire for or against a certain kind of commodity. It is not difficult to understand that the purchase of a new coat, if such coat were not necessary, would entail an appeal of a very opposite character than were such necessity established.

The results indicate in Problem II very definitely the fact that price appeal is the least important factor where necessity has already been established and that merchandising and institutional advertisements have about an equal amount of influence under such circumstances. This information is highly significant in planning seasonal advertising campaigns when certain merchandise is absolutely necessary to meet these seasonal requirements. The trade is familiar with conditions as they exist, as to when the original purchaser buys or when the individual buys who is not too much concerned with the necessity for a coat as the hope of

obtaining a bargain or keeping abreast of the styles. So we know definitely that the person forced because of need of a new garment, price appeal is the least important thing, provided the other factors or money, etc., do not function, and in order to check against this, Problems V and VI were introduced which, in the last analysis, are a check against the validity of these findings.

The figures for Problem II follow:

PROBLEM II

		Male		*Female*		*Total*	
	1.	93		142		235	
A	2.	81		122		203	
	3.	62	236	86	350	148	586
	1.	94		150		244	
B	2.	39		72		111	
	3.	103	236	130	352	233	588
	1.	49		60		109	
C	2.	116		158		274	
	3.	71	236	136	354	207	590

PROBLEM III

To answer Problem III, you must assume the following conditions:

You *now have a coat* and are only *vaguely interested* in a new one, THEREFORE,

You are *not* decided on a *particular kind* of new coat NOR have you determined the *amount* you may spend for the new coat NOR have you determined *where* you may purchase a new coat

In view of these assumptions, answer the following question:

Which of the three advertisements furnished you (A, B, C) in their rank order would most arouse your interest and be most likely to crystallize your vague interest, causing you to purchase a new fur coat?

1st choice ——— 2nd choice ——— 3rd choice ———

This problem introduces the element of the particular kind of coat, amount of money to be spent and place of purchase, the motivation or desire for a new coat more or less at a minimum. Since the interest has not been established in the direction of a new coat, the appeal must be of a different sort. In this situation the institutional advertisement is the least forceful, while that of price appeal and merchandising seem to be pretty evenly divided as factors in selection.

Figures for Problem III follow:

PROBLEM III

		Male		*Female*		*Total*	
	1.	93		145		238	
A	2.	99		127		226	
	3.	44	236	80	352	124	588
	1.	61		99		160	
B	2.	39		80		119	
	3.	136	236	173	352	309	588
	1.	82		109		191	
C	2.	98		145		243	
	3.	56	236	99	353	155	589

PROBLEM IV

Problem IV reads as follows:

To answer Problem IV, you must assume the following condition:

> You *have definitely* decided to purchase a fur coat. In view of this decision, which of the three advertisements (A, B, C) would exercise the greatest influence in causing you to make the purchase at the Mason & Adams Store?
>
> 1st choice ——— 2nd choice ——— 3rd choice ———

Many prospective buyers find themselves in the situation presented by Problem IV. A decision has been made to pur-

chase but money, place of purchase, and specific kind of merchandise have not definitely been established. Thus it is interesting to determine whether or not the introduction of one or more of these other factors is as important as it would be to stress a proposition in its entirety. The conclusions which follow from Problem IV are a good evaluation of the respective advertisements, when a limited number of factors are involved. Under these conditions merchandising and institutional advertisements rank higher than price appeal.

RESULTS

PROBLEM IV

		Male		*Female*		*Total*	
A	1.	89		143		232	
	2.	89		129		218	
	3.	58	236	79	351	137	587
B	1.	96		143		239	
	2.	46		69		115	
	3.	94	236	139	351	233	587
C	1.	51		65		116	
	2.	101		153		254	
	3.	84	236	134	352	218	588

PROBLEM V

Problem V reads as follows:

To answer Problem V, you must assume the following conditions:

You *have decided* to purchase a fur coat

You have decided your *finances will not permit* you to spend in excess of *$150* (cash or installment plan)

In view of these assumptions, which of the three advertisements furnished you (A, B, C) in their rank order, would exercise the greatest influence in causing you to make the purchase at the Mason & Adams Store?

1st choice ——— 2nd choice ——— 3rd choice ———

Problem V is concerned with one major decision, since it has been determined that the individual is to purchase the fur coat and has only a limited amount of money at his disposal. Quantitatively, the results are satisfactory in the degree of correlation, because under such circumstances the institution should play but a small rôle in selection. The institutional advertisement is the least convincing and of least importance in directing the buyer to a particular establishment, while on the other hand there is a fair degree of evidence which shows merchandising to be the most important and price appeal only a runner-up directing an individual to a particular place.

Figures for Problem V follow:

PROBLEM V

		Male		*Female*		*Total*	
	1.	135		197		332	
A	2.	98		148		246	
	3.	3	236	8	353	11	589
	1.	0		6		6	
B	2.	12		22		34	
	3.	224	236	324	352	548	588
	1.	101		150		251	
C	2.	126		182		308	
	3.	9	236	21	353	30	589

PROBLEM VI

Problem VI reads as follows:

(Note: Do not permit your judgment in Problem V to prejudice your answer in this problem.)

To answer Problem VI you must assume the following conditions:

You *have decided* to purchase a fur coat

You have decided your *finances will not permit* you to spend in excess of *$500* (cash or installment plan)

In view of these assumptions, which of the three advertisements furnished you (A, B, C) in their rank order, would exercise the greatest influence in causing you to make the purchase at the Mason & Adams Store?

1st choice ——— 2nd choice ——— 3rd choice ———

This problem is similar to Problem V, excepting that the expenditure involves five hundred dollars. This means several things. It means, first, that the amount of risk at issue is much greater and that reputation and character of the organization to be dealt with is a greater issue. The evidence supports a pre-analytical "guess" that the institutional appeal is the greatest factor in the decision, merchandising second, and price appeal third. This certainly establishes the necessity for the occasional appearance of the institutional advertisement, especially in advertising expensive merchandise. The institutional advertisement must be used in an entirely different way, the display and body of the advertisement less conspicuous and altogether presenting a more dignified, higher type of message than either the merchandising or price appeal advertisements.

RESULTS

PROBLEM VI

		Male		*Female*		*Total*	
	1.	59		90		149	
A	2.	137		221		358	
	3.	40	236	42	353	82	589
	1.	164		249		413	
B	2.	48		69		117	
	3.	24	236	36	354	60	590
	1.	13		14		27	
C	2.	51		64		115	
	3.	172	236	274	352	446	588

PROBLEM VII

Study carefully the three advertisements (A, B, C) and considering yourself as a consumer, i.e., one who buys fur coats, what is the rank order of the effectiveness (appeal, persuasiveness, compelling force) of these advertisements in causing you to purchase your next coat at the Mason & Adams Store?

1st choice ——— 2nd choice ——— 3rd choice ———

State briefly the reasons for your selection:

The reason for my first choice is.....

.....

The reason for my second choice is.....

.....

The reason for my third choice is.....

.....

Problem VII establishes the effect of purpose. The absence of amount of money, reasons for selection or necessity for purchase, elevates the decisions in this case to that of a purely academic decision. It is a statement of preference, of analytic discrimination of abstract, intellectual judgments. Under these circumstances isolating a type of judgment, and eliminating material values, the institutional type of advertisement has the greatest appeal, persuasiveness and compelling force, both for males and females.

RESULTS

PROBLEM VII

		Male		*Female*		*Total*	
A	1.	75		118		193	
	2.	86		129		215	
	3.	73	234	104	351	177	585
B	1.	119		176		295	
	2.	48		67		115	
	3.	67	234	107	350	174	584
C	1.	40		58		98	
	2.	100		154		254	
	3.	94	234	138	350	232	584

PROBLEM VIII

If you were looking through a newspaper and there appeared in different issues of the publication, advertisements A, B, C, what is the rank order in which they would arrest your attention?

1st choice ——— 2nd choice ——— 3rd choice ———

Attentional factors and the persuasiveness of the original appeal are the prime considerations of this problem. Because of the speed with which the average newspaper is read and the fact that in the major newspapers, the greatest comfort is not maintained if the paper is held fully spread, it is important to determine which advertisement in a given moment of attention seems to carry with it the greatest force. The institutional advertisement is first, merchandising second, and price appeal third. This, to us, considering the fact that the advertisements were differently sized, which in a sense vitiates the value of this particular problem in itself, certainly should present to those interested in advertising as a science a very fertile field for investigation, as to whether size in itself is not a matter of prime importance.

RESULTS

PROBLEM VIII

		Male		*Female*		*Total*	
	1.	69		131		200	
A	2.	83		109		192	
	3.	80	232	112	352	192	584
	1.	96		137		233	
B	2.	46		93		139	
	3.	90	232	122	352	212	584
	1.	67		150		217	
C	2.	103		84		187	
	3.	62	232	118	352	180	584

PROBLEM IX

On general principles, as the advertisements appear in their entirety, what is the order in which you "like" them?
1st choice —— 2nd choice —— 3rd choice ——

Problem IX is not unlike Problem VII, except that there is introduced the factor of fundamental appeal by the employment of the word "like" in quotations. There is a certain æsthetic compulsion which looms important by virtue of the wording of the question. The institutional advertisement again assumes first place; merchandising, second; and price appeal, third.

RESULTS

PROBLEM IX

		Male		*Female*		*Total*	
	1.	38		65		103	
A	2.	84		146		230	
	3.	112	234	142	353	254	587
	1.	165		247		412	
B	2.	37		59		96	
	3.	32	234	47	353	79	587
	1.	31		41		72	
C	2.	113		148		261	
	3.	90	234	164	353	254	587

Problem X reads as follows:
In a brief statement or phrase, what is in your opinion the fundamental message conveyed by the advertisements? Use space designated for each advertisement.
In my opinion the message conveyed by advertisement "A" is .....
In my opinion the message conveyed by advertisement "B" is .....
In my opinion the message conveyed by advertisement "C" is .....

This problem is primarily concerned with an attempt to establish whether or not the analysis of the lay reader is in agreement with the advertising expert in designating advertisement "A," "Merchandising"; advertisement "B," "Institutional" and advertisement "C," "Price Appeal."

There follow herewith representative answers to Problem X in which there is readily observable a unanimity of interpretation.

ADVERTISEMENT "A"

From "Female" Questionnaires

Sale of a variety of sample coats at reasonable prices.
There is a coat at every price.
The greatest number and selection of coats at prices to suit everyone.
Great variety of models at great variety of prices.
A sale of sample furs at about half price.
That Mason & Adams have fur coats at all prices.
A price and style to suit everyone.
Good large stock of coats and bargains.
Sample furs of every type at half price for the buyer who wants a smart fur coat.
Sample furs—great variety.
All kinds of coats at whatever price you want to pay.
Wide range and explicit information.
Exact information as to price and styles.
That there is in progress a sale of fur coats varying in prices and styles.
Variety in kind of coat and vast range of prices.

From "Male" Questionnaires

Large selection at conservative prices.
Great quantity of sample furs at all price ranges.
Good quality at low price.
Pick what you want; we have it at fair price and easy terms of payment.
Just a sale for advertisement purposes, a slight reduction.
Stock and price.

Wide field of choice, style as well as reasonableness.
A variety of choice in coats and prices.
Wide assortment.
Shows styles, range of prices, trade in old furs, and small deposit.
An appeal to all classes of buyers, to stimulate buying.
Quality and quantity at low price.
Great variety in type of coat and cost.

ADVERTISEMENT "B"

From "Female" Questionnaires

Good coats are at Mason & Adams for those who can afford them.
Furs—Mason & Adams—Exclusive.
That the firm has exclusive styles.
To be exclusive buy your fur coat here.
That the store deals in exclusive fur fashions.
That Mason & Adams is an exclusive shop for discriminating women.
That Mason & Adams is an exclusive shop.
The security which is assured the consumer who buys at such a reliable firm.
An anniversary sale which will carry on the honor and trust of America's largest retail fur establishment by its great values.
The exclusiveness of the store.
We are catering to exclusive trade.
Quality and name of the firm.
Quality backed by a good name.
Quality backed by reliability.
That Mason & Adams are carriers of exclusive fur fashions.

From "Male" Questionnaires

Smartness and exclusiveness of Mason & Adams furs at reasonable cost.
A store that gets out such an ad should have stylish and refined products for sale.
The quality of the store's furs.

An exclusive shop which invites your patronage.
Quality and distinctive coats sold by a quality store.
Mason & Adams is an exclusive establishment.
High class goods at a reliable and exclusive firm.
We are a very exclusive shop.
The chance to go to a reputable house and get quality, though you must pay for it.

ADVERTISEMENT "C"

From "Female" Questionnaires

A one day clearance of coats at less than wholesale cost.
That coats are to be sold at a bargain which are of good value.
That the firm carries lower priced coats, the less expensive ones.
Large price reduction.
An opportunity to get a bargain.
The great drop of price in fur coats at the sale.
Price appeals almost entirely.
One day sale of limited supply of fur coats at less than wholesale.
Great reductions in low priced coats.
Appeal to the pocketbook.
Price or amount of money involved.
Coats have been reduced in price.
An appeal to the bargain hunter.
A sale of all types of fur coats at reduced prices.
Wonderful bargain sale.

From "Male" Questionnaires

Price. Great reductions.
Appeal to bargain hunters.
Buy now and at rock bottom prices.
Here are coats reduced before new year begins.
The fact that coats are being sold at a bargain price.
Should appeal to a person of taste but limited means.
Here is an opportunity for bargain hunters.
Cheapness.
One day sale of furs at greatly reduced prices.
A chance to get a fur coat cheap.

SUMMARY AND CONCLUSION

Psychological analysis of advertising is at present receiving much attention with especial emphasis on the questionnaire method. For the guidance of those interested the present study offers a specific contribution:

1. The questionnaire must be devised to insure:
   a. Proper selection of material
   b. Effective type of interrogation
   c. Correct interpretation
   d. Control of variables
2. Some attempt should be made that only interested persons should participate in furnishing data.
3. The competence of the individual is a factor of importance.
4. Effective advertising to be economically sound must appeal alike to both male and female readers regardless of the sex of the ultimate purchaser.
5. The time element should receive serious attention.
6. All three types, "Merchandising," "Institutional," and "Price Appeal" should be employed. Our findings indicate the following:
   a. Academic interest is aroused by merchandising type because being highly descriptive, it allows for the functioning of verbal processes.
   b. Where purchase has been decided or where money is not the prime consideration the institutional type of advertising is most effective.
   c. With limited money, merchandising and price appeal are most necessary.
7. Seasonal advertisers should stress merchandising in terms of need, money and value.
8. Advertising of price alone rarely if ever justifies its regular and frequent employment.
9. The efficacy of the pure type of advertisement, kinds of type, legend, æsthetic compulsion, the amount of reading material, size, are definitely determinable and the findings can be profitably employed.
10. There is true value in having specific purpose in advertising

as indicated by the analyses of results of the problems herein presented in which the variables are controlled. Psychological or other analyses are of no value unless control is exercised. The process may not differ from that observed in any other experimental problem. Despite the obvious limitations of the Method of the Questionnaire, the method here employed makes for valid results. While authentic and interesting data for the advertiser are to be found in the answers to the problems emphasizing the value of specific purpose, in the present test T-N-25 in which the determination of purpose or motivation is effected by the conditions of the questions, there is established the value of the method only when employed under controlled conditions.

THE VALIDITY OF CERTAIN QUESTIONS WHICH PURPORT TO MEASURE NEUROTIC TENDENCIES*

CARNEY LANDIS and S. E. KATZ

New York State Psychiatric Institute and Hospital

THE Bernreuter Personality Inventory Blank (2) consists of 125 questions, each of which may be answered by "yes," "no" or "?." By a method of differential evaluation the extent to which the response to a single question is symptomatic of each of several personality traits may be determined. The four traits which this test is designed to estimate or evaluate are, 1) Neurotic tendency (designated as B1-N), 2) Self-sufficiency (designated as B2-S), 3) Introversion-extroversion (B3-I), and 4) Dominance-submission (B4-D).

The neurotic tendency scale (B1-N) particularly interested us, since it seemed that the test might be validated by comparison with an external criterion through its application to a neurotic or psychotic population. Bernreuter's (4) statement concerning this scale is as follows. "The individual who scores high on the B1-N scale shows a tendency toward a neurotic condition. Such an individual often feels miserable, is sensitive to blame, is troubled by useless thoughts, by shyness, and by feelings of inferiority. He feels shut off from other people, he frequently day-dreams, and worries both over things that have happened and over things that may happen. The person who scores low on the B1-N scale is an emotionally stable person. He is rarely troubled by moods, by worries, or by the criticism of others. He is self-confident, and is a doer rather than a day-dreamer." Bernreuter (2) also states that those scoring above the 98 percentile for neurotic tendency would benefit by psychiatric or medical advice.

The present study has been directed toward the following points:

*The data on which this paper is based were collected in part by Miss Joan Henning and Mr. Louis DeAngelis. This paper was presented in part before the American Association for the Advancement of Science, Boston, December 1933.

1. Our interest in the general validity of such an inventory when used as a measuring or diagnostic instrument with neurotic and psychotic patients.

2. Might such an inventory be of use in differentiating the various clinical psychiatric diagnostic groupings?

3. The investigation, wherever possible, of the reasons for discrepancies between the answers given by the patient when taking the test, and the known past record of that individual as shown by the case history, with respect to the points covered by the question.

METHODS

The Bernreuter test was administered either in small groups or individually, to 184 house patients and to 40 out-patients of the Psychiatric Institute. The distribution of those patients by diagnostic grouping was as follows: Manic-Depressive (Manic), 21, Manic-Depressive (Depressed), 27, Dementia Præcox, 64, Psychoneurosis, 61, and Others (General Paralysis, Psychopathic Personality, etc.), 51. All of our group were "voluntary" patients at the Institute. They were all fairly cooperative and interested in the test.

RESULTS

The percentage distribution of the scores arranged by percentile

TABLE I.

Percentile Score	Total Group N.	Total Group Per Cent	M-D. Manic N.	M-D. Manic Per Cent	M-D. Dep. N.	M-D. Dep. Per Cent	D. P. N.	D. P. Per Cent	Psychoneurosis N.	Psychoneurosis Per Cent
0- 10	22	10	3	14	3	11	4	6	2	3
10- 20	14	6	3	14	0	0	5	8	1	2
20- 30	14	6	2	10	2	7	4	6	1	2
30- 40	19	8	1	5	1	4	6	9	3	5
40- 50	16	7	2	10	2	7	4	6	4	7
50- 60	16	7	3	14	2	7	5	8	4	7
60- 70	25	11	5	24	4	15	5	8	6	10
70- 80	20	9	2	10	4	15	5	8	7	12
80- 90	22	10	0	0	3	12	11	17	9	15
90-100	56	25	0	0	6	21	15	23	24	38
	224		21		27		64		61	

score for the total group and for each of the diagnostic groups for B1-N is shown in Table I.

The following points are brought out by Table I.

1. Of the entire group of 224 psychotic and psychoneurotic patients, only 56 (25 per cent) scored above the 90th percentile while 22 (10 per cent) of the patients scored below the 10th percentile.

2. Of the dementia præcox patients, 23 per cent scored above the 90th percentile and 6 per cent scored below the 10th percentile.

3. None of the manic-depressive, manic patients, scored above the 80th percentile and 14 per cent scored below the 10th percentile. Of the manic-depressive, depressed, 21 per cent scored above the 90th percentile and 11 per cent scored below the 10th percentile.

4. Of the psychoneurotic patients, 39 per cent scored above the 90th percentile and 3 per cent scored below the 10th percentile.

5. Although there was some variation in the scoring between the various diagnostic groupings (the psychoneurotic group differing significantly in its median score from the other groups), it is quite obvious that the test falls short as a means of determining neurotic tendency as far as the individual psychiatric patient is concerned.

Why does the inventory give such unsatisfactory results when used with neurotic and psychotic patients? With this question in mind, the written case histories of 90 of the 224 patients who had taken this test were carefully studied. An attempt was made to compare the personality traits of the patient, as described by the relatives or as stated by the medical and hospital clinical history, with the trait described in the test. It is obvious that all of those traits given in the inventory did not have a counterpart in every clinical history. In the comparison of the case histories with the inventory test statements, certain stock phrases which frequently appear in case reports were evaluated and the following equations made: "Stubborn, insistent on having his own way" was equated with "Trying to get your own way even if you have to fight for it" (Question 11). "Sensitive, easily of-

fended" was equated with "Feelings easily hurt" (Question 51). "Fault-finding" was equated with "Critical of other people" (Question 116). "Social gathering" (Question 28) was taken to mean any occasion when 2 or 3 gathered together for social purposes. "Depression" was considered "Just miserable" (Question 8). Similar interpretations were given to the other bits of phraseology in the anamnesis in order to make them comparable to the test material wherever possible.

The time element conveyed in the question presented another difficulty. Questions containing the word "ever" were comparatively easy to evaluate because they obviously refer to any part of the subject's life. Other questions seemed to refer to reactions, usually emotional, which vary a great deal with circumstances and can be answered unequivocally only according to the recent past. Such items as, "Do you often feel just miserable?" (Question 8), were, therefore, compared with behavior extending over an arbitrarily selected period of 3 months just prior to the test. In Question 50, "Do you usually try to avoid arguments?" the term "usual" was compared with the subject's behavior during the year previous to the test. Statements as to what the patient had been like more than a year previous to the test, were not used because of the large amount of change which might have taken place since that observation was made, particularly since our patients were practically all "acute" cases having their first episode of mental disease.

The comparison of the answers given on the test and the relevant material obtained from the case history has been expressed as a percentage of the number of agreements. The mean per cent of agreement for various B1-N percentile intervals was as follows: B1-N score 0-10, 15 cases, mean agreement score 50 per cent; 10-20, 8 cases, 60 per cent; 20-50, 21 cases, 70 per cent; 50-80, 24 cases, 73 per cent; 80-90, 8 cases, 79 per cent; and 90-100, 14 cases, 91 per cent.

The median per cent agreement for these 90 patients was 73. An examination of the B1-N scores of those patients having less than 73 per cent agreement compared with those having more than 73 per cent showed that the distributions were slightly skewed in opposite directions. Of the patients above the median

of 73 per cent, 31 per cent have a neurotic percentile score of 90 or over and only 4 per cent score below 10. For the patients below the median of agreement, none scored above the 90th percentile while 29 per cent scored below the 10th percentile. For the 18 patients with an agreement score of 90 per cent or above, none had a B1-N score below 30 and 8 patients (44 per cent) scored in the 90th percentile or above.

The emotional mood of the patient at the time of the test was compared to the agreement score, since, theoretically, elated patients might be overly optimistic in their answers concerning their character traits. We found that patients making low agreement scores showed no predominating tendency for either an elated or depressed mood. However, 16 out of 20 patients who were depressed at the time of the test were in the upper quartile of agreement scores.

Very little or no relationship was found between the agreement score and intelligence, age, grade reached in school, insight (as judged by the physician in charge of the case), psychiatric diagnosis or prognosis. Analysis of the source of the statements in the history with respect to their agreement with the test score showed a slight tendency for the recorded statements of the patient to agree with the test more often than did the statements made by the relatives concerning the patient. This tendency was not statistically significant.

In an effort to get at the psychological determinants of the disagreement between the test and the history, 30 of the patients, 9 men and 21 women, were interviewed. All of these, except 4 women, had a B1-N score below the 25th percentile or above the 75th percentile. In the interview, the patient was shown his answer to a disagreeing question and told to explain more fully what he had meant by it. The patient, in his explanation, often volunteered reasons for the disagreement. When he did not give what seemed to be a cogent reason, he was asked explicitly if he did not exhibit this trait or behavior at some recent time.

This, of course, gave him an opportunity to state that he had recovered from this symptom and so may account for a number of statements of this nature. Unfortunately there seemed no

other way to obtain the information without using such a leading question.

The reasons for disagreement were classified in the following way:

1. *Carelessness.* The patient changed his answer and said that the earlier one had been due to "hurry," "confusion," etc.

2. *Different definition.* For example, the experimenters had interpreted Question 62 ("Do you make new friends easily?"), as "Do you get acquainted with people easily?" For 2 patients the explanation was, "Yes, I get acquainted easily. But real friends I'm choosy about." This is an example of a justifiable difference of interpretation. Other interpretations seemed less reasonable and looked rather like hedging. For instance, an answer to Question 14 ("Do you consider yourself a rather nervous person?") was, "I'm nervous but not *above average.*"

3. *Different frequency.* Here the experimenter had estimated that a certain symptom occurred less frequently or more frequently than the patient thought.

4. *Recovery.* The patient claimed recent recovery from the symptom.

5. *Contradiction.* Here the patient simply denied the statement in the history and could give no reason for the disagreement. Occasionally the patient's behavior in the interview seemed to contradict his statement. For example, an elated manic patient stated that she was not excited.

6. *No definite answer* given. In some cases the patient did not

TABLE II

Reason for disagreement	B1-N Score 0-25	B1-N Score 25-75	B1-N Score 75-100
Carelessness	0	3	3
Definition	4	1	2
Different frequency	3	0	4
Recovery	8	1	0
Contradiction	27	1	7
No definite answer	6	1	2
Total No. questions	48	7	18

really explain his answer to the question but rambled around, making no very definite statement and sometimes contradicting himself.

The reasons for disagreement, arranged according to the B1-N percentile score, are given in Table II.

From Table II we may conclude: 1) Contradiction, in the sense of lack of insight or of a different impression which the patient holds in opposition to the impression of his associates, is the most frequent source of a lack of agreement between past history records and test results. This factor operates markedly in those individuals who make a low B1-N score. It constitutes the basis for 56 per cent of the disagreements in the lowest quartile scores, and is the basis for 37 per cent of the disagreements in all ascertained instances. 2) Disagreements due to carelessness appeared in the upper three quartiles but not in the lowest. This may mean that individuals in the lowest quartile were covering up their disagreements with a "rational" explanation wherever possible. 3) The patients in the upper B1-N quartile of scores, seemed to have more valid reasons for their test statements, which disagreed with their history, than did those individuals in the lowest quartile.

In 63 specific instances questions on the test were answered with a question mark when the case history gave definite evidence for a "yes" or "no" answer. According to the history, 53 of these should have been answered in a neurotic way, and 10 in a normal way. Forty-three of these "?" answers were given by patients with an agreement score below the mean. The history indicated that 33 out of these 43 answers should have had a neurotic answer and 10 a normal answer. According to the history the twenty "?" answers of subjects with an agreement score above the mean should all have been neurotic answers.

THE VALIDITY OF CERTAIN WEIGHTED TRAITS

An analysis of all the relevant data which we have at hand on the 10 items which Bernreuter (4) considered to be of the greatest importance in the diagnosing or describing of "neurotic tendency" is given in Table III. To this list we have added 2 questions (14 and 2), on which we had particularly relevant data. Table

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Bernreuter No.	Bernreuter questions	Neurotic response	Positive diagnostic value	Negative diagnostic value	Weighted score B1-N	Per cent agreement when answered "yes"	Per cent agreement when answered "no"	No. of "yes" answers checked	No. of "no" answers checked	Per cent of normals answering this question neurotically	Per cent of schizophrenics answering this question neurotically	Per cent of M-D psychotics answering this question neurotically	No. of "well-adjusted" (Group I—Thurstone) answering neurotically	No. of "neurotics" (Group II—Thurstone) answering neurotically
8	Do you often feel just miserable?	y	30+	−30+	7	92	85	37	26	...	...	...	0	40
51	Are your feelings easily hurt?	y	30+	−30+	7	100	19	25	21	...	...	...	2	37
73	Does some particularly useless thought keep coming into your mind to bother you?	y	30+	−30+	7	100	40	7	5	...	...	...	0	35
104	Do your feelings alternate between happiness and sadness without apparent reason?	y	30+	−30+	7	73	50	15	20	43	25	32	5	41
24	Are you troubled with shyness?	y	30	−30+	6	100	38	16	13	28	35	26	0	33
39	Do you worry too long over humiliating experiences?	y	25	−30+	5	83	83	6	6	55	46	39	1	10
72	Are you troubled with feelings of inferiority?	y	30	−26	6	95	50	21	12	48	36	35	0	36
106	Do you worry over possible misfortunes?	y	30	−22	6	92	53	35	19	35	46	35	2	37
5	Can you stand criticism without feeling hurt?	n	21	−26	5	33*	100*	3	4	24	38	29	0	25
54	Do you often feel lonesome when you are with other people?	y	20	−22	4	100*	...*	1	0	21	32	25	1	42
14	Do you consider yourself a rather nervous person?	y	..	..	3	100	0	27	19	..	..	..	0	32
2	Do you day-dream frequently?	y	..	..	5	90	25	11	12	43	31	25	3	40

TABLE III

Comparison of validation of 12 questions in the Bernreuter test.

Columns 2, 3, 4 and 5 are from Bernreuter (4) Table VIII. Column 6 is from the Bernreuter (2) B1-N scoring key. Columns 7, 8, 9 and 10 give the agreement scores determined in this study. Columns 11, 12 and 13 give the percentage frequency occurrence in normal, schizoid and manic-depressives, from Page, Landis and Katz (6). Columns 14 and 15 are from Thurstone and Thurstone (7) Table I.

* Figures are meaningless for lack of sufficient data.

1 Thurstone No.	2 Bernreuter No.	3 "Schizoid" questions which are significantly more "normal" than "psychotic"	4 Per cent of normals answering Yes	5 Per cent of schizoids answering Yes	6 Per cent of M-D psychotics answering Yes	7 Bernreuter weighting for Yes answer	8 No. of "well-adjusted" (Thurstone) answering Yes	9 No. of "neurotics" (Thurstone) answering Yes
4b- 9	116	Are critical of others	69	39	29	3	4	32
2b-40	17?	Think it possible for other people to influence their actions	58	39	34	3	0	8
3a-12?	91	Are often occupied with their own thoughts	84	70	65	3	0	32
1b- 3	56	Are not tactful	47	33	29	−1	14	1
2b-10	2	Day-dream a lot	43	31	25	5	3	40
2a-16	25?	Have but little faith in human nature	35	22	20	2	11	2
4b-21	93	Find it difficult to make decisions	48	34	32	3	7	29
4b-4	72	Have feelings of inferiority	48	36	35	6	0	36

TABLE IV

Questions more frequently normal than psychotic (Page, Landis and Katz (6). Columns 1 and 2 give the number of the item, or its nearest equivalent in meaning, as it appears in Thurstone and Thurstone (7) and in Bernreuter (2). Columns 3, 4, 5 and 6 are from Page, Landis and Katz (6) and give the percentage frequency of the occurrence of the neurotic answer in normals, schizophrenics and M-D psychotics. Column 7 is the Bernreuter (2) B1-N weighting for the neurotic answer. Columns 7 and 8 are from the Thurstone and Thurstone (7) Table I giving the standards appearing for the "internal consistency" validating groups.

IV gives a comparative analysis of 8 items which Page, Landis and Katz (6) found to be significantly "normal" more often than "psychotic."

The following points are brought out by Tables III and IV.

1) "Yes" answers, to B1-N questions which are most diagnostic of neurotic tendency, tend to have a high agreement with outside validating criteria.

2) When the answers given by psychotic patients (schizophrenics and manic-depressives) are compared to those given by a normal control group it is apparent that questions 104, 39, 72 and 2 are normal, significantly more frequently than they are psychotic. Question 5 tends to be schizophrenic more often than normal.

3) Taking 8 items which have been shown to indicate normality (Table IV) we see that the Bernreuter weighting places a neurotic value of 2 to 6 points on each item, save one, 56. In other words, answering these 8 questions in what is really a "normal" fashion would entitle one to 24 positive points on the B1-N, neurotic scale.

4) Thurstone's method of internal consistency "validates" as neurotic every item in Tables III and IV except Thurstone No. 2b-40. The discrepancy between the methods of validation will be discussed later in this paper.

5) This analysis which is based on 20 selected items from the 125 in the Bernreuter scale presents a fair case for the entire scale so far as the items are diagnostic of neurotic tendency. We have analyzed other items similarly and find essentially the same results.

THE QUESTION OF VALIDATION

Bernreuter (4) validated his inventory principally on the basis of its high correlation with such other tests as the Thurstone (7) Neurotic Inventory, the Bernreuter (3) Self-Sufficiency Test, the Laird (5) C2 Introversion Test and the Allport (1) Ascendance-Submission Reaction Study. These tests gave coefficients of correlation with the Bernreuter scoring which range from +0.67 to 1.00. The Thurstone (7) Neurotic Inventory, on which Bernreuter based his claim for the validity for the B1-N scale,

was validated by the "criterion of internal consistency." Briefly this criterion was established as follows. A list of 223 questions, most of which are descriptive of traits supposedly associated with the neurotic personality, was given to 694 college freshmen. The authors of the list tentatively decided whether "yes" or "no" constituted the neurotic answer to each question. All blanks were then scored and the total number of neurotic answers, based on the tentative scoring key, was determined for each subject. The 50 highest scores and 50 lowest scores were then sorted from the entire group and the number of times each of the 50 high or 50 low scoring subjects agreed or disagreed with the tentative answer to each item, was determined. If a majority of the high scorers agreed or a majority of low scorers disagreed with the tentative answer, that question was considered significant and the tentative answer verified. If this verification did not occur the question was discarded or the tentative answer reversed. Such a procedure really sets up an arbitrary "personality trait cluster" which is self-sufficient and self-verified and need have no relation to external validating criteria save the fact that the same word, in this case "neurotic," is used to describe a self-verified personality cluster and a variety of abnormal personality of interest to the psychopathologist.

However, since the word "neurotic" is used in both cases, there is a tendency to identify the two concepts without further proof. Bernreuter's statement (2) that those scoring above the 98th percentile on the B1-N scale would probably benefit from psychiatric advice, indicates that Bernreuter has thought of the B1-N "neurotics" as similar to, if not identical with, psychopathological "neurotics." Thurstone's (7) likewise implies an identification of these two concepts, as they state, "If we succeed in finding many of the neurotic personalities among the freshmen for psychiatric consultation or supervision, it will be something gained" (7, page 15). Since this identification of concepts has become common it is well to see if the use of the external validating criteria gives results which are comparable to Thurstone's method of internal consistency.

We have applied Thurstone's method of determining internal consistency to our data, using 125 psychotic cases and 125 normal

individuals who had been matched with the psychotics with respect to age, social status and education. This group of 250 are entirely different from the previous groups mentioned in this paper. From this new group we have taken the 18 highest B1-N scores and the 18 lowest scores, since this number is approximately the same per cent of our group that 50 is of Thurstone's group. Of the 18 high ("neurotic") scorers 9 were normal individuals and 9 were neurotic. Of the low ("well-adjusted") scorers 6 were normal and 12 were neurotic. This is essentially a corroboration of Table I which showed that the neurotic population tended to approximate the distribution of the normal population except for a certain amount of skewing at the upper end of the distribution.

It can be readily shown, as Zubin (8) has done, that the application of the method of internal consistency to questionnaire-like material made up of nonsense syllables would give a resulting high "validity" to some of the items in the test. In other words, the method sets up a series of self-verified truisms, with respect to the items of the test.

CONCLUSIONS

Thurstone and Bernreuter have provided the psychologist with an instrument which was said to differentiate the emotionally stable from the emotionally unstable, the mentally well-adjusted from the neurotic. In actual practice it has been variously reported that those receiving high scores are neurotic. Our study confirms this. However, the converse is not true. Those whose scores indicate that they are emotionally stable are not necessarily so. It seems probable that the admitting of those traits which constitute the neurotic scoring on this scale may justify the diagnosis of neurotic personality, but that these individuals should be judged neurotic on no other basis than this test, and so advised to seek a psychiatrist, has not been demonstrated. Conversely, a low score on the test does not mean that there is no necessity for the individual to seek medical aid and assistance.

With respect to the validity of the questions, we have shown that high scorers are frequently neurotic in the psychiatric sense of the term. In any event they seem to be telling the truth about

themselves, whatever may be the nature of their neurosis, psychosis, or "normality." Low scorers in our neurotic group obtain low scores for a variety of reasons, chief of which seems to be that of flat contradiction of their personal opinion of themselves with their past record of behavior. It would be exceedingly interesting to attempt such validation on a normal group, particularly investigating the low scorers to see whether they also would have such high disagreement scores as we have found, and whether such scores are based on contradiction. We would hazard the prediction that a similar study to ours but based on normal individuals with a similar validation of the truth of the items would give essentially the same distribution of agreement scores.

To summarize, the personality inventory gives results which are indicative of poorly adjusted personality when the individual receives a high score; when the individual receives a low score, we are not justified in drawing any conclusion concerning the satisfactoriness of his adjustment to life. On direct validation we found that about three-fourths of the self-descriptive statements given by neurotic individuals really agree with the externally ascertained facts concerning those individuals.

There is a need for a questionnaire type of scale which is based not on preconceived notions of what does or does not constitute undesirable personality traits, but which will make use of the analysis of many questions which may or may not seem to be immediately relevant, and which give, when appropriately treated, results showing that these traits do tend to occur together in variously widely separated groups.

SUMMARY

An analysis of certain of the questions which are included in the Bernreuter Personality Inventory and in the Thurstone Neurotic Inventory with respect to the validity and diagnostic weightings, has been made. It has been shown that approximately three-fourths of the answers which are given by neurotic patients are verified by their previous history. The greatest source of disagreement lies in a contradiction between the patient's opinion of himself and his past record. As one should expect the disagreements are more prevalent in the case of those neurotic

individuals who make low scores on these tests; however, the prevalence is not as great as it should theoretically have been. The scores on the tests are not adequate measures of neurotic tendency, so far as psychiatric patients are concerned. Validation by means of the method of internal consistency has been shown to be inadequate. The personality inventory or schedule gives results which are indicative of poorly adjusted personality when a high score is received. If the individual receives a low score, no conclusions concerning his emotional adjustments to life may be drawn.

REFERENCES

1. Allport, G. W., A test for ascendance-submission. J. Abn. & Soc. Psychol., 1928, *23*, 118-136.

2. Bernreuter, R. G., The personality inventory. Stanford Univer. Press. Stanford Univer., Calif., 1931.

3. _____ The measurement of self-sufficiency. J. Abn. & Soc. Psychol., 1933, *28*, 291-300.

4. _____ The theory and construction of the personality inventory. J. Soc. Psychol., 1933, *4*, 387-405.

5. Laird, D. A., Detecting abnormal behavior. J. Abn. & Soc. Psychol., 1925, *20*, 128-141.

6. Page, J., Landis, C., & Katz, S. E., Schizophrenic traits in the functional psychoses and in normal individuals. Amer. J. Psychiat., 1934, *13*, 1213-1225.

7. Thurstone, L. L. & Thurstone, T. G., A neurotic inventory. J. Soc. Psychol., 1930, *1*, 3-30.

8. Zubin, J., The method of internal consistency. J. Educat. Psychol., 1934, *25*, 345-356.

A STUDY OF LITERARY APPRECIATION

M. E. BROOM

State Teachers College, San Diego, California

THE study of literature in the schools has been advocated on the grounds that literature is one of the desirable outlets for the emotions.[1] Elsewhere the writer has presented information on the controversy which exists as to the nature of esthetic literary experience.[2] The intermediate position on the question appears to recognize literary appreciation as the product of intellectual understanding, coupled with such emotional response as results from the active understanding on the part of the reader. Judd writes: "Where there is no ability to understand, there is a lack of appropriate response; where the response is absent or inappropriate, the resulting emotion will be deficient or perverted."[3]

The present report discusses an attempt to study the emotional content of esthetic literary appreciation. A number of literary passages were selected by a group of English teachers who were cooperating with the writer, from which five were selected finally for inclusion in a pencil and paper test.[4]

---

[1]P. Sandiford, *Educational Psychology*. New York: Longmans, Green and Company, 1929. Pages 139-140.

[2]M. E. Broom, "Literature and Esthetics." *High School Teacher*, 8:293-294, October, 1932.

[3]C. H. Judd, *Psychology of Secondary Education*. Boston: Ginn and Company, 1927. Page 219.

[4]The passages were from the following sources:

Conrad, Joseph, *The Rescue*. Garden City, New York: Doubleday, Page and Company, 1920. Pages 288-289. (From "The boat swam forward on the glare . . . (to) . . . suggesting assurance and power.")

Hewlett, Maurice, *Life and Death of Richard Yea-and-Nay*. New York: Chas. Scribner's Sons, 1919. Pages 253-254. (From "The Tower of Flies . . . (to) . . . in the thick air. . . . All dead things . . . (to) . . . down there in the ooze.")

In each of these five selections the judges indicated a certain number of words and phrases, the understanding of which they felt would involve the subject in an emotional reaction. The choice was finally limited to the fifty most emotional words, and these were arranged in multiple response form in the test, with three alternate words, with each one of which there was something logically or psychologically wrong.

Presland's "A London Idyll," for example, presents a foggy London night, with two young lovers standing hand in hand, whispering, in the drizzling rain. No other living thing is about save a stray cat slinking across the wet pavement. If our three alternate words with whispering in the text of the poem are conversing, chattering, and conferring, obviously there is something wrong logically or psychologically with each.

The multiple response test was prepared in mimeographed form for administration to high school and college students. Several hundred subjects at these school levels responded to this first pencil and paper test. The resulting scores were correlated with scores yielded by the Pressey X-O Test for investigating the Emotions, Form A; the Laird Personal Inventory, C-2; and other similar measuring instruments. The following tabular material includes typical correlation findings.

Later investigation of the reactions yielded during subsequent administrations of the pencil and paper test and its modified forms indicated that this technique was little likely to yield results of value for solving the problem. The next attempt to find an approach yielding adequate insight into the problem involved the use of the psychogalvanometer. Measurements were taken of the psychogalvanic reflex reactions of subjects during the oral reading of four of the five passages listed in footnote 4 (Presland's "A London Idyll" was eliminated, leaving all prose selections),

---

Lawrence, D. H., *Sons and Lovers.* New York: Boni and Liveright, 1922. (Modern Library, No. 109.) Pages 351-352. (From "There was a great . . . (to) . . . under a darkness of leaves.")

Cather, Willa, *My Antonia.* Boston: Houghton Mifflin Company, 1918. Pages 223-224. (From "At last . . . (to) . . . toward the tent of themselves.")

Presland, John, "A London Idyll." From J. M. Bachelor and G. Greever's *The Soul of the City.* Boston: Houghton Mifflin Company, 1923. Pages 195-196. (From "A heavy sky . . . (to) . . . hand in hand.")

TABLE I

Test correlated with the objective emotions—reading test.[5]	Number of subjects	r	School grade of subjects
Terman Group Test of Mental Ability (mental age)	62	.48	9-12
Average judgments of three English teachers as to pupil-literary appreciation[6]	62	.69	9-12
Van Wagenen Reading Scale, English Literature Alpha	62	.65	9-12
Kennon Literary Vocabulary Test (Form B)	62	.59	9-12
Thorndike Intelligence Examination for College Entrants, total equivalent scores	115	.25	14-college graduates
Thorndike Int. Test, Part III, (reading comprehension)	115	.39	14-college graduates
Pressey X-O Test for investigating the emotions, Form A, i. d. score	115	.13	14-college graduates
Pressey X-O Test ... t. aff. score	115	.26	14-college graduates
Laird Personal Inventory, C-2	84	.14	14-16

together with three new selections.[7]

[5]The reliability of this test was positive .87 (45 cases, college students.)

[6]The reliability of the average judgments of three English teachers was positive .76 (63 cases, high school students.)

[7]Hardy, Thomas, "The Three Strangers." In *The Great Modern English Stories,* edited by E. J. O'Brien. New York: Boni and Liveright, 1919. Pages 20-21. (From "In the meantime the company . . . (to) . . . on the side of the law.")

Post, M. D., "The Great Legend." In *Representative American Short Stories,* edited by A. Jessup. Page 935. (From the beginning of the story to "the bronze opposite the Madeleine?"

Whitman, S. F., "That Famous Love Affair." *In The World's One Hundred Best Short Stories,* edited by Grant Overton. New York: Funk and Wagnalls Company, 1927. Volume 7, Pages 179-180. (From "He drove his heel . . . (to) with a broken neck.")

The psychogalvanic technique is well known. Wheeler[8] and Landis[9] give good short discussions of it. In the present investigation the following instruments were used: a hand electrode constructed by the experimenter, slightly modified from that described by Ruckmick,[10] a Leeds and Northrup Type "S" Testing Set (including a self-contained 4½ volt battery, a self-contained galvanometer, and a five dial Wheatstone Bridge),[11] and a D'Arsonval galvanometer (Central Scientific Company, Catalog F, No. F4445.) The D'Arsonval galvanometer was used instead of the less sensitive galvanometer included in the Testing Set.

The nature of the responses yielded by different subjects to the fiction material can be shown best by reproducing one passage, indicating by underlining the specific words to which some definite electrical disturbances began. What takes place probably is this: Visceral discharges take place, of which the subject is unaware where certain (stimuli) words are encountered during the reading of the passage, which result in a change in the electric conductivity of the body. (Even though at times the reaction "built up" over following words or phrases following the complete quotations, the underlined words are grouped together, the number letter indicating the subject, and the numbers indicating the amount of galvanometric reflection.) Only major disturbances for each subject have been considered. The passage is from Post's "The Great Legend:"

"It did not occur to us to sleep. We sat by the tiny fire made of bleached bones and some handfuls of grasses burned by the *sun* and as hard as parchment.

"We were in the desert south of the Zar'ez. In the vast, motion-

---

[8]Wheeler, R. H. (Editor) *Readings in Psychology.* New York: Thomas Y. Crowell Company, 1930. Page 199. The Editor's Note introducing Nancy Bayley's "A Study of Fear."

[9]Landis, C. "The Expressions of Emotion." In *The Foundations of Experimental Psychology,* edited by C. Murchison. Worcester, Mass.: Clark University Press, 1929. Pages 514-516.

[10]Ruckmick, C. A. "A New Electrode for the Hathaway Galvanic Reflex Apparatus." *American Journal of Psychology,* 42: 106-107. January, 1930.

[11]Central Scientific Company (Chicago, Illinois,) Catalog F, No. F4750.

less ocean of sand; billows of sand like a disturbed sea paralyzed by a stroke of sorcery. The moon filled the *desert,* the mountains behind us, the waves of *sand,* the salt bed of the Zar'ez shimmering like a silver *disk*.

"It was all a weird, saffron, devil's *color*.

"The Arabs were asleep about us, each muffled in his *burnoose* like a mammoth goat *fleece*. De Morney sat beyond me across the tiny fire. Now and then far out beyond *us* in this remote, limitless ocean of sand, there came the rolling of a drum, the roll of innumerable, distant *drums* blended into a faint, indistinguishable, menacing war *sound*.

"It was a night for any fantastic notion. And the spirit of the great war that had just closed was on all of us.

" 'The drums of Imam,' he said, and he extended his arm toward the sleeping Arabs. 'Every nation has its great legend; its heroic *spirit* to *awaken* in the hour of need.'

"He paused. Then he added: 'Is it a *dream,* or does the thing happen?'

"He was a big, heavy, strange man: French crossed *deeply* with the *Oriental*. He belonged in the mystery of this desert. He was the authority of France in this nomad world. Everybody had heard about him. He held Algeria to the Republic through every wile of Envar Pasha and his innumerable mad *mullahs*. Nothing could move *him*. When you looked at him you thought of a figure in *basalt*. He sat now on a rug in this desert, his feet under him, his face inscrutable, like an *Oriental*.

" 'There is *truth* in the thing,' he said, 'strange, inconceivable *truth* in it. Do you know the bronze opposite the Madeleine?"

In some instances the deflection appeared to be of the startle variety, quickly generating and just as quickly dissipating. The response of subject V, herself an Oriental, to the word "Oriental" was of this sort. Other responses, while generating at times slowly and at times quickly with a large deflection, either maintained the large deflection with very slow and gradual decrease or even increased the amount of the deflection for a period of time before the beginning of the decrease. Subject R's response to the word "dream" was of this type. Apparently the more stable response, such as subject R's to "dream," indicated more

WORD	SUBJECT	DEFLECTION
sun	E	2.0
desert	P	2.3
sand	O	1.2
disk	E	2.0
color	N	8.5
burnoose	J	2.6
fleece	F	6.4
beyond (us)	O	4.2
drums	(H	3.1
	(A	1.8
	(G	2.2
	(D	4.8
sound	(T	2.3
	(I	1.6
(weird) thing	S	15.2
spirit	L	3.8
awaken	K	2.3
dream	(R	11.2
	(Q	6.0
	(U	7.2
deeply	B	2.3
Oriental	V	5.5
mullahs	M	5.3
(Move) him	(K	3.6
	(P	3.7
	(D	12.2
basalt	S	4.5
Oriental	V	6.2
truth (appears twice in passage)	L	2.3

nearly a true emotional condition aroused by the stimulus word. Even in such instances it is difficult to determine whether the word itself was associated with something not in the context which caused the emotional disturbance, or whether the emotional disturbance was direct result of the fiction context itself.

As a check on the validity of the psychogalvanic technique in this experiment, ten subjects were asked to read three non-fiction prose selections, giving psychogalvanic reflex records during the

period of reading.[12] The reactions were found to be quite uniform in character, all small in amount (no one over 0.2 scale divisions, and only some four or five over 0.1 scale division), and all probably within the range of normal disturbance during the non-emotional situations.

Both qualitative and quantitative differences have been detected by the galvanometer among the individuals who responded to the passages used in this study, but it appears that little has been done to solve the problem of the nature of æsthetic experience. Probably the galvanometer itself is not an adequate instrument for use in a study of this type, since the period of the galvanometer in reaching a "true" reading is slower than the duration of the galvanic wave which is measured.[13] Too, there is some question as to whether the psychogalvanometric reflex is specific to emotion.[14] There is an interesting field for research here to the person interested in the theoretical problem of the nature of æsthetic experience, and also to the educationist because of its probable practical results bearing on the curricular content of literature courses in the secondary schools.

---

[12]The passages included the following:

The second paragraph in the Preface to the Oxford Test of *The Comedies of Shakespeare*, edited by W. J. Craig. Pages xxxvii-xxxviii. London: Oxford University Press, 1927.

The paragraph on Wundt's Tridimensional Theory in R. H. Wheeler's *The Science of Psychology*. Page 176. New York: Thomas Y. Crowell Company, 1929.

The paragraph on the Constitution of the Confederacy in John Fiske's *The Beginnings of New England*. Page 158. Boston: Houghton, Mifflin Company, 1889.

[13]R. C. Davis, "Factors Affecting the Psychogalvanic Technique." *Archives of Psychology*, No. 115, page 9. New York: Department of Psychology, Columbia University, 1930.

[14]C. Landis, loc. cit., page 514.

See also, G. Murphy, *Historical Introduction to Modern Psychology*. New York: Harcourt, Brace and Company, 1930. Page 374.

THE RELATIONSHIP OF THE NUMBER OF QUESTIONS MISSED ON THE OTIS MENTAL TESTS AND THE ABILITY TO HANDLE OFFICE DETAIL

S. N. STEVENS and E. F. WONDERLIC

Northwestern University

THE number of questions missed and skipped on the Otis Mental Ability Test and its relation to certain work habits of a group of branch office managers in a large personal finance company suggests further use of intelligence testing in industry. The Otis Self-Administering Test of Mental Ability, higher examination, Form A, was administered, using a thirty-minute time interval. The number of questions missed and skipped was the sum of those incorrectly answered and those omitted prior to the last question attempted. It is assumed that the last question answered is the last question attempted by the subject.

The number of questions missed and skipped was then compared to the work habits of branch office managers which involved their ability to organize work and systematize the detail operations of their individual offices. A manager's ability to handle this detail work was judged from the accumulated reports of the supervisory staff. These supervisory reports, which concern the men's abilities, development and limitations, are made following visits which occur at six-week intervals and collect in his personnel portfolio. The criticisms are unsolicited observations and at no point was there a definite request made for reports on a manager's ability to handle office detail.

From a careful inspection of the complete personal histories of 160 managers, we have been able to isolate those individuals who, at any time during their service as managers, have been criticized for their method of systematizing the office, handling detail, and generally following the routine work. The individuals studied have served the company as branch office managers one or more

years and each of them has had at least two additional years' experience as assistant in various branch offices of the organization.

Examples of the typical comments which were used as an indication of ability to handle office detail are presented below with the number of questions missed by the individuals on the Otis Test. Cases A, B, C, and D are from GROUP A, individuals who have been criticized. Cases V, W, X, Y, and Z are managers in GROUP B, individuals who never have been criticized for laxness in handling detail.

CASE A: Attempted 69 questions
Missed 31 questions

"It is almost impossible for him to arrange his work systematically."

Ten months later:

"His one weakness is handling of detail and in getting others to do the same."

One year later:

"Still very weak in detail but progress is being made in correcting this."

CASE B: Attempted 74 questions
Missed 29 questions

"His routine work is very poorly arranged. He did not watch his employees and results obtained. Has been inclined to let things go. Office work has been behind."

CASE C: Attempted 64 questions
Missed 16 questions

"This man is aggressive and ambitious but rather impatient and sometimes careless."

CASE D: Attempted 54 questions
Missed 2 questions

"I had quite a talk with him regarding the lack of organization in his office. *He has the ability* to organize his work and force if he would only spend the effort to do so."

CASE V: Attempted 64 questions
Missed 1 question

"The girls made several errors last month. He was greatly concerned about this. He takes great pride in having everything just right."

CASE W: Attempted 49 questions
Missed 1 question

"Goes into details very thoroughly and writes a thorough reference application."

CASE X: Attempted 75 questions
Missed 4 questions

"Very thorough and competent. The office functions with exceptional smoothness and work is done quickly and quietly."

CASE Y: Attempted 75 questions
Missed 3 questions

"At the time of being transferred and installed at the new office, he had the force together and explained to them how he wanted the window sheet, report sheet, etc., properly marked. He insists that all detail work be properly done."

CASE Z: Attempted 59 questions
Missed 4 questions

"He is very prompt, very industrious, good on details and persistent to a marked degree."

The above examples are only a few of the typical illustrations of instances having been isolated when managers are criticized or commended for handling the office detail organization. Many other comments of like nature may be found in their work histories as well as in the histories of the other individuals studied above.

A study of such complete work records of 160 managers, then, makes it possible to group them according to their ability to handle office detail and organization. Table I, Group A, consists of those who, from time to time during their work performance as managers, were severely criticized in regard to their ability.

Group B consists of those managers who have never been criticized because of their lack of ability to organize and follow up details. In Group A there are 62 managers, 12 of these, or 19.4 per cent of this group, missed and skipped less than 16 questions. Fifty, or 80.6 per cent, missed more than 15 questions. The average number of questions missed and skipped for Group A is 22.56. In Group B there are 98 managers, and 73, or 74.5 per cent, missed and skipped less than 16 questions. Twenty-five, or 25.5 per cent of this group, missed and skipped more than 15 questions. The average number of questions missed and skipped by this group, who were never criticized, was 11.96.

These facts suggest that those managers who missed 15 or more questions are very likely to be criticized for their ability to

DISTRIBUTION OF NUMBER OF QUESTIONS MISSED AND SKIPPED IN THE OTIS TEST BY BRANCH OFFICE MANAGERS

In groups according to their ability to handle office detail and organization—160 cases

No. Questions Missed-Skipped	GROUP A Mgrs. Severely Criticized			GROUP B Mgrs. Never Criticized		
	No.	Accum. Freq.	Per Cent of Group	No.	Accum. Freq.	Per Cent of Group
0 thru 5	2	2		18	18	
6 " 10	3	5	19.4	26	44	74.5
11 " 15	7	12		29	73	
16 " 20	19	31		12	85	
21 " 25	9	40		8	93	
26 " 30	10	50	80.6	2	95	25.5
31 " 35	7	57		2	97	
36 " 40	2	59		1	98	
41 " 45	3	62		0	98	
Group Averages		22.56			11.96	

handle office details. Apparently this item measures a mental habit which is detrimental to accurate thinking and it is extremely difficult for individuals who miss a considerable number of these questions to organize their thoughts and actions. Under pressure, we have found that this type of individual is very likely to become

impulsive rather than use sound problem-solving methods. A further study of the number of questions missed and skipped is now in progress in an attempt to establish norms for adult performance.

The findings of this study suggest, then, that the number of questions missed and skipped on the Otis Mental Ability Test may have real diagnostic significance in any work situation where men are being employed for tasks which involve careful handling of office details under pressure.

THE PSYCHO-EDUCATIONAL CLINIC

PAUL A. WITTY and VIOLA THEMAN

Northwestern University

IN 1914, Wallin found that 16 universities, 3 normal schools, and 7 medical schools had psychological or psycho-educational clinics.[1] Theodate Smith in 1914 described eight clinics connected with institutions of higher education.[2] Few descriptive accounts or reports dealing exclusively with the psycho-educational clinic have appeared in recent years. Witty,[3] in 1930, and Donald Snedden,[4] in 1933, outlined practises or made recommendations for certain psycho-educational or mental hygiene clinics. During 1932-33, Town and others summarized the results of a questionnaire study dealing primarily with the training and activities of clinical psychologists.[5] Although that part of the investigation concerned with the psycho-educational clinic was incidental, several important facts regarding the nature of courses which have been designed for training psychologists were set forth. There appears to have been no thorough descriptive account of psycho-educational clinics since Wallin's report was published in 1914.

I *A Questionnaire Study of Psycho-Educational Clinics (1932-33)*

In 1932-33, the writers studied the organization and practises

[1] Wallin, J. E. W., *The Mental Health of the School Child.* Yale University Press, New Haven, 1914, p. 25 and p. 57.

[2] Smith, T. L., The development of psychological clinics in the United States, *Pedagogical Seminary,* 21: 143-153, March, 1914.

[3] Witty, Paul A., Some results secured in a psycho-educational clinic, *Journal of Applied Psychology,* 14: 160-178, April, 1930.

[4] Snedden, D., The work of a psycho-educational and mental hygiene clinic, *Journal of Educational Sociology,* 6: 516-524, May, 1933.

[5] Town, C. H., Merrill, M. A., and Brown, A. W., Report on the survey of the training and duties of clinical psychologists, *The Psychological Exchange,* 2: 109-114, August, September, 1933.

of psycho-educational clinics affiliated with institutions of higher learning. Psycho-educational clinics were considered to include all clinics in which students are trained to examine and to report upon the educational attainment as well as upon the mental development of the cases.

Information was sought from 380 universities, colleges, teachers' colleges and normal schools whose catalogues implied that psycho-educational-clinic service was available. Of the 340 that replied, only 85 stated that they maintained clinics of the psycho-educational type. To these 85, detailed questionnaires were sent; these were returned by 58. Several reported clinics whose major functions are not pertinent to this report (e.g., personnel work with college students, speech clinic, psychopathic clinic, and so forth). Of the schools which did not return the questionnaire, four are thought to operate psycho-educational clinics (Stanford, Columbia, New York University, and The University of Pennsylvania).

The following summaries deal exclusively with the data received from the 34 schools which provide not only educational and psychological service for school children but which also train students to administer and interpret tests. Although one can make only an estimate of the completeness of this survey one may guess that about 85 per cent of the psycho-educational clinics in institutions of higher education located in the United States are herein represented.[6]

A. *Organization of the Clinics*

(1) Establishment. These clinics have been in operation from 1 to 25 years; the Training School at Vineland, N. J., is the oldest, having been established in 1907. The median number of years of existence is 4.

(2) Personnel. The number of staff members ranges from 1 to 6, and the median is 2. Ninety per cent of the clinics have at least 1 member who holds a Ph.D. degree, and 8 members (in 5 clinics) have M.D. degrees. Sixty-five per cent of the clinics have at least one staff member who be-

[6] The names of the clinics will be sent to any one interested in having them.

longs to the American Psychological Association, and fifty-one per cent of the clinics are represented in the American Association for the Advancement of Science.

(3) The main purposes of the clinics follow in rank-order of frequency.
   (a) Providing service to schools, social agencies, parents, and others, through administering diagnostic tests and suggesting remedial materials and methods to be used in effecting educational, vocational, or social adjustment.
   (b) Training students in administering and interpreting test results.
   (c) Research:
      (1) Study of deviates.
      (2) Causes of and remedies for learning difficulties.
      (3) Experimental work with remedial materials.

(4) Types of cases accepted.
Cases of the following classifications are tested in the various clinics:

	Per cent of clinics examining each type
Preschool	81
Primary	96
Elementary	96
Junior high school	96
Senior high school	90
Adult	69

Eighteen schools limit the cases tested to the public school, university training schools, and welfare agencies; three reject mentally subnormal cases, while two deal with amentia cases only. Three-fourths administer tests in places outside the university or the school laboratory. However, most of the work is carried out in the laboratory. Very few schools charge fees for testing. One school charges $.50 to $1.00; another charges $25.00 for each case; and a few charge occasional fees. Sixty-five per cent exact no charge.

B. *Classification of Cases*

(a) The number of cases tested during one calendar year ranges from 15 in one clinic to several thousand in another, with a median of about 150. More boys than girls are tested, the sex ratio being 4:3.

(b) The range in chronological age is from 6 months to old age. Only five clinics test children during the first year of life, and 4 clinics examine more adults than children.

(c) Classification of IQ's indicates a distribution departing considerably from that for an unselected population. Individual clinics, of course, show much variation in this respect. The average IQ of the cases reported by 20 clinics is 83; the standard deviation is 21.

C. *The Course in the Study of Clinical Procedures*

The clinic course is offered with about equal frequency in the department of education or of psychology. Fifty per cent of the clinics admit college juniors as students, and 83 per cent enroll seniors. However, seniors and graduate students comprise most of the enrollment, which is limited numerically. The number of students ranges from 4 to 40, with a median of 12.

(1) *Prerequisites.* All schools list introductory courses in Education and in Psychology as prerequisites for the clinic training. Some require little more, while others prescribe many related subjects. Teaching experience is a prerequisite in only 3 clinics.

(2) *Number of Courses and Credit.* Credit received varies with the nature of the course. Four schools offer more than one course in clinical work. The Ohio State University has a curriculum rather than a course, consisting of 3 required laboratory courses stressing test techniques for which 6 to 9 quarter hours of credit are given. (The class meets twice weekly for two-hour sessions.) These courses are followed by four quarters of "Psychological Clinic Practicum." During the first two quarters the student acts as "student clinician"; the last two quarters are generally devoted to field work in the children's hospital or in the juvenile court.

The Training School at Vineland is a special case: "The Clinic is a psycho-educational one and the actual examining is, for the

most part, carried on by fellow students who do what might be considered a clinical interneship here. . . . An affiliation with Rutgers University enables the student to get college credit for his work."

The University of Nebraska provides a series of courses which total 20 semester hours of credit; the Universities of Iowa and Kansas provide for 12 semester hours in several related courses.

Credit ranging from 2 to 6 semester hours is usually granted; the median is 3. More than half of the schools offer the clinic course during one semester only.

D. *Clinical Activities Engaged in by Students*

The student's work varies according to his stage of development in using the techniques of clinical procedure. The following list indicates the rank-order of activities on the basis of time allotment:

Administration of tests, attendance at lectures, round table discussions, observation of demonstration testing, presentation of oral reports. In addition, nine other activities were mentioned in single instances.

E. *Supervision of Student Testing*

Although the directors of 3 clinics report no supervision of student testing, the majority provide very close and almost continuous supervision as well as careful scrutiny of the interpretative accounts of the tests.

Only 2 schools permit the students to read the directions for the Binet-Simon Revision and other tests. Sixty per cent require memorization, and the remainder permit occasional reference to a book of directions. Nevertheless, all directors seem to feel that it is essential that the student have rather complete verbal mastery of the testing techniques in order that he may observe the child for qualitative evidences of adjustment or its lack.

The measures of the student's achievement include evaluation of: general proficiency with many tests, accuracy in administering tests and determining results, keenness of observation of incidental factors occurring during the testing, ability to secure and maintain "rapport," and capacity to suggest appropriate and successful remedial measures. Emphasis is placed by eighteen directors upon the student's ability to evaluate qualitative aspects of the

child's behavior during the testing (a clinician's point of view rather than that of a psychometrist). Ability to test cases representing all chronological ages is stressed by two directors. In the majority of clinics the decision regarding student proficiency and success rests almost entirely upon the instructor's subjective judgment.

II. CLINICAL EXAMINATIONS AND CASE REPORTS

Significant variations are found among the clinics in the types and combinations of data which are collected as well as in the final recording and interpreting of the examinations.

A. *Types of data collected for the complete case study*

(1) Mental tests are administered in all the clinics. Some revision of the Binet-Simon test, usually the Stanford, is always administered. About one-third of the clinics train students to give more than one revision; one clinic uses the Hayes-Irwin Revision with the blind. Frequently reported and in the following order of frequency, were the Porteus, the Pintner-Paterson, and Grace Arthur Scales.

(2) Educational tests are given in all clinics. Most frequently cited was the Stanford Achievement Test. Diagnostic tests were mentioned by less than one-third of the respondents. It appears that the clinicians are interested usually in ascertaining the general educational level of the case rather than his detailed educational status.

(3) Character tests are used in most clinics. Those most frequently named were the Thurstone, the Bernreuter, and the May-Hartshorne series.

(4) Special aptitude tests, administered in more than 4 clinics, include Seashore's Measures of Musical Talent and the Stenquist Mechanical Aptitude Test. The Kent-Rosanoff Association Test was listed by four of the respondents.

(5) A rather complete social history of each case is obtained (frequently from the referring agency).

(6) Physical examination results are usually secured by 81 per cent of the clinics. Cooperative agencies are used for

obtaining complete physical diagnosis, while tests of vision, hearing, height and weight are made in many of the clinics.

(7) Psychiatric service is alleged to be provided by 54 per cent of the clinics; this is usually obtained through affiliation with other agencies. This service appears to be neither far-reaching nor well-organized.

(8) Eighty-one per cent of the clinics report follow-up service in varying amounts and degrees of thoroughness. The usual practises include reexamination of cases, and provision for an orderly, checked therapeutic program.

B. *Supervision and form of the case reports*

In 54 per cent of the clinics the original draft of the case report is written by the student who is held responsible for scoring and interpreting the tests. The report is then reviewed and edited by a staff member. In these cases the general form for writing the reports is outlined by the instructor. A few exceptions follow: the report is written by the student and the staff member in conference; a staff member writes the report after he has consulted the student; a student writes the report as he would a practise theme, and a revision is not made. The cooperative work of student and instructor, and the editing of student accounts are the two methods used most frequently.

Printed or mimeographed forms for writing the case reports, particularly for recording data concerning the test results are employed by only 11 schools. A wide variation is found in the form of these data sheets. They are generally one page in length, containing spaces for inserting test results and norm equivalents. Additional descriptive reports are usually sent to the person or agency which refers the case to the clinic. These are about 500 words in length usually, and consist of a summary statement and recommendations for particular cases.

C. *Conclusions and Recommendations*

One is impressed by the wide variations which exist among the clinics in regard to almost every item concerning the organization as well as the teaching techniques.

Very few institutions of higher education offer a curriculum and provide opportunity for clinical work which appear adequate to prepare individuals for mental hygiene and psychoclinical work. In 1914 Wallin outlined an extremely exacting course of training for the specialist in this field, one that would require three or four years beyond the B.S. degree. It included a study of anatomy and pathology, of public and personal hygiene, of the common physical defects, of nervous and mental diseases, of psychopathology and psychotherapy, and of pediatrics. He advocated that students have teaching experience in the elementary school as well as in training schools for teachers; he included also the study of all phases of corrective pedagogics as well as a thorough foundation in the differential pedagogy which applies to the particular types of cases the student expects to handle.

In no case did the respondents indicate provisions as inclusive as those advocated by Wallin. The Ohio State University reported the most complete organization for the training of clinicians. One must bear in mind, however, that many of the schools have not, as yet, undertaken the certification of workers in this particular field. (Ohio State University was the only one in which this was mentioned.) Usually the courses are offered for teachers of special classes as well as for those who plan to specialize in clinical psychology and pedagogy.

The student who plans to specialize in psycho-educational work will do well to choose his school with extreme care, considering carefully the amount and type of opportunities provided by the institution. He will not find all the courses advocated by Wallin in any one school but he will find a majority of them provided in the set-up of a very few.

PART II

A REPORT OF 318 CASES STUDIED IN THE PSYCHO-EDUCATIONAL CLINIC AT NORTHWESTERN UNIVERSITY, 1930-1933

A. *Organization of the clinic*

The psycho-educational clinic in the School of Education of Northwestern University was established in September, 1930.

It has a two-fold function: (1) to instruct a limited number of graduate students in administering tests and in interpreting results; (2) to provide educational diagnosis and psychological service for children referred to the clinic by agencies located in the neighboring communities.

The staff of the clinic consists of the director and his two assistants.

All the school grades from the kindergarten to the university are represented in the cases sent to the clinic. Any of the following problems may be referred: educational, grade placement, guidance, conduct, mental defect, and mental abnormalities in the way of special aptitude or disability. Children are generally sent by the neighboring schools or by parents, although cases come occasionally from other departments of the University.

The children are usually tested in the clinic laboratory at the University. However, when a study is made of a particular school group or class, the testing is done at the place which is most convenient and appropriate. Occasionally testing is carried out in a school in which a special class is organized.

The clinic equipment includes the materials for administering the Stanford-Binet, the Porteus Test, the Performance Scales, and diagnostic and standardized educational tests. Small rooms are provided for individual testing. Apparatus is available for testing vision, hearing, height, and weight.

Because of the limitation in facilities, only a small number of graduate students are permitted to enroll in the clinic course. The number usually does not exceed eight. Specialized courses are not always made prerequisite, but candidates are rigidly selected according to academic records and the director's judgment of their probability of success in this endeavor. In the majority of cases the students have had experience as teachers or as administrators. In addition, they present evidence of course work or its equivalent in abnormal psychology.

No laboratory fee is charged for students enrolled in the course, nor are fees exacted from those who are tested. The clinic course proceeds from September to June, and students may enroll for one or for both semesters. The amount of credit received upon completion of the course varies from two to four hours for each

semester. Students meet once a week for a two-hour period. They spend an additional unscheduled period each week in the clinic. Their activities include: observation of demonstration testing, attendance at lectures, administration of tests, and presentation of oral and written reports of case studies. The first two activities are those which most frequently characterize the initial weeks of the semester.

Student testing is observed each week by the members of the staff. The length of time spent by a staff member in observation depends upon the proficiency of the tester and the difficulties involved in diagnosing particular cases. Errors made by the student are usually corrected immediately by the instructor, but care is exercised in the method of making the correction so that "rapport" will not need to be reestablished between the examiner and the child. Each student spends in addition to the two-hour period of testing, an hour or more with the director and his assistants in analyzing and revising the reports.

Emphasis is placed on the establishment and maintenance of "rapport" during the testing. This is facilitated by requiring students to show complete verbal mastery of the directions for giving the tests. Thus, freedom from reference to books and other pamphlets enables the examiner to direct his attention to the child's behavior and to the quality of responses. Throughout the course effort is put forth to use a "clinical" approach rather than the psychometrist's method.

B. *Procedure in the clinic*

The technique used in making the clinical examinations is similar to that reported by Witty in describing the clinic operated at the University of Kansas.

The following types of data are secured during the period of observation: (a) physical, (b) mental, (c) educational, (d) emotional.

(1) *Physical.* Vision and hearing are examined and the condition is recorded. The customary method of scoring, recording, and interpreting the condition of the eyes is employed; i.e., 20/20 vision is normal. Auditory acuity is gauged by the whisper test as well as by the audiometer. Height and weight are measured

and the results are compared with the Baldwin-Wood standards. During the time of making these physical tests, the examiner notes and records any unusual condition of the skin or postural irregularity.

The purpose of this part of the observation is to discern physical limitations which may handicap the child in his school work. Conspicuous weaknesses or anomalies are referred to physicians for more complete diagnosis and treatment.

The examiner is zealous to identify sensory defects which might obliterate or distort the imagery of the child, and to provide corrective measures where poor health or disease diminishes the potentialities of the child for success in the work of the school.

(2) *Mental.* To ascertain the mental status of the child, the Stanford and other revisions of the Binet-Simon Test of Intelligence are employed. The Porteus Test of Motor Coordination, Prudence, and Forethought is given to supplement the Binet results. All children who have language handicaps, speech defects, or defective vision or hearing are examined by appropriate performance scales. The intelligence test results are not considered final estimates of the child's ability, but merely as points of departure for further study of the child.

(3) *Educational attainment.* All children are given educational tests when specific inabilities are reported or observed. For this purpose, the following tests are used most frequently:

1. Stanford Achievement Tests
2. Gray Oral Reading Paragraphs
3. Gates Silent Reading Tests
4. Compass Diagnostic Tests in Arithmetic
5. Public School Publishing Company Achievement Tests

Many of the reliable diagnostic tests are used from time to time. In addition, the child's proficiency in school work is estimated by observing his proficiency with regular school texts and materials, and by consulting his school history. One extremely important phase of the work deserves mention. Because of the interest of the director of the clinic in gifted and talented children, many children are brought to the clinic who are thought to be gifted intellectually, or especially talented. For some of these children rather complete data, comparable to those secured by

L. M. Terman in his investigations of gifted children, are collated. For others, who appear especially talented, appropriate measures are used. For example, one boy (IQ 156+), who is unusually talented in music has been given the Seashore, the Kwalwasser-Dykema, the Kwalwasser-Ruch and the Diemer tests in music. His progress has been observed and recorded over a period of five years.

(4) *Nervous stability.* Although a course relating to the nature and needs of non-typical children is not requisite in the clinic training, most students in the clinic are taking such a course concurrently with their clinic work. In this course emphasis is placed upon understanding and treating the unstable child whose difficulties are occasioned largely by malfunctioning of mental ability rather than by a deviation in the quantitative amount of ability. However, the major symptoms of mental instability are reviewed and therapeutic measures are examined critically in the clinic course.

During the various examinations, the student identifies and records any conspicuous tendencies toward instability which are displayed by the child. These tendencies are revealed by the symptomatic responses occurring during the examinations and by the general behavior of the child. The following are examples of such symptoms: range upon Binet, bizarreness of response, tendency toward perseveration, flightiness, automatic or sound reactions, general nervousness, excessive giggling, and so forth.[7] The recreation interview is profitably used in this phase of the work.

C. *Reporting Results*

A simple clinic form report is used in recording numerical expressions of the four types of data described above. The items concerning the social history which are obtained from the school or from the parent in informal visits are added.

These data are collated by the student and organized in a case report which follows a prescribed organization: general statement,

[7] See Witty, Paul A. Some results secured in a psycho-educational clinic. *Op. cit.* pp. 161-164.

physical status, mental test results, educational status, observation cues, and finally, recommendations.

A complete folder containing tests and interpretation is submitted to a member of the staff; it is read and alterations and suggestions for improvement are incorporated. Then, it is returned to the student in order that he may make the changes or justify his own interpretation. Finally, the complete folder is filed in the clinic after the corrections have been made.

These reports in whole or in part (particularly the sections on recommendations for remedial work) are sent to the persons who have sought help from the clinic. Retests are made only when the director feels that a retest will be of value to the person tested, of interest to the members of the clinic, or when unusual or undesirable conditions have prevailed during the first testing. "Follow-up" work is not provided for all the cases. Nevertheless, several groups of children have been organized in neighborhood schools in which the remedial suggestions have been carried out and in which progress has been checked. One report of this work has already been published.[8] Numerous individual remedial programs have been planned, progress has been followed, and gains have been reported.

D. *Data concerning the cases tested*

The following data concerning verbal and motor intelligence are treated for two groups: the cases reported from September 1930 to June 1932 are in one, and those reported from September 1932 to June 1933 in another.

(1) *Mental test results.* The 154 cases tested during the first two years include seventy-three girls and eighty-one boys. The median IQ of the girls is 106.2, (mean 106) S.D.=20.2; and of the boys, 97.8 (mean 101.6) S.D.=20.4. The median IQ of the whole group is 101.4 (mean 103.79) S.D.=20.4.

The 164 cases tested during the third year of the clinic include eighty-five girls and seventy-nine boys. The median IQ of the girls is 110.2 (mean 111.3), S.D.=21.5; and of the boys, 108.6 (mean 109.48) S.D.=22.1.

[8] Witty, P. A., and LaBrant, L. L. Some results in remedial instruction in reading, *Educational Trends*, 2: 7-13, January, 1933.

By comparing these data with those of Terman's 1000 unselected children for whom he found the mean IQ-100, S.D.=13, one notes that the clinic groups are noticeably superior and decidedly more variable. The cases include 3.7 per cent potentially feeble-minded; 19.1 per cent very superior; and 7.5 per cent gifted.

The following tabulation sets forth the range in IQ and the range in C.A.:

Year	Range in IQ	Range in C.A.
1930-32	54 to 159	3 to 18 years
1932-33	41 to 166	3 to 19 years

The coefficients of correlation between the mental ages obtained from the Porteus test and those from the Stanford-Binet are given below:

Year	r — Girls	r — Boys
1930-32	+.642 P.E. ±.06	+.756 P.E. ±.04
1932-33	+.817 P.E. ±.04	+.883 P.E. ±.02

A twofold study was made of the occupational status of the parents of the children tested: (1) to determine the percentage of cases of extremely high or low IQ in each of four occupational groups classified according to the United States Census Reports; (2) to ascertain the coefficient of correlation between the IQ as determined by the score on the Binet test and the P.E. value assigned on the Barr Scale to the occupations of the fathers.

For these two studies the data accumulated during the three years were combined. The case studies of twenty children of IQ 135 or above and of nineteen of IQ 85 or below contained statements of the occupations of the fathers. The occupations are divided into four classifications: professional, commercial, industrial, and public service. Each has sub-heads by which the classification is further delimited. The percentages of cases in each classification are indicated in Table I.

There are twice as many fathers of children whose IQ is 85 or below in the industrial group as in the professional group,

TABLE I

Number and Per Cent of Cases of Extreme Intelligence Quotient (IQ 85 or Below and IQ 135 and Above) in Each of Four Occupational Groups

Occupational Group	Cases Tested at Northwestern U. Clinic				Terman's Results (1)
	IQ 85 or Below		IQ 135 or Above		IQ 140 and Above
	Cases	Per cent	Cases	Per cent	Per cent
Professional	4	21.0	10	50.0	29.1
Commercial	6	31.5	6	30.0	46.2
Industrial	8	42.0	4	20.0	20.2
Public Service	1	5.2	0	0.0	4.5

while two and one-half times as many fathers of children of IQ 135 and above are classified in the professional group as in the industrial group. Terman's[9] results in his study of the occupations of the fathers of gifted children do not indicate so marked a concentration as this, although the data accumulated by Witty[10] show a similar tendency for the occupations of the father of gifted children to fall in the professional group. Although the proportions vary, all studies point toward the same general tendency. They support clearly the contention of L. M. Terman regarding the concentration of the children of high IQ in the homes of professional people.

The coefficient of correlation between the IQ and the P.E. value of the father's occupation on the Barr Scale was $+.35$ P.E. $\pm.056$; (this is lower than the coefficient reported by Witty: $+.48$ P.E. $\pm.02$).[11]

(2) *Educational status and grade placement*

The children tested in the clinic represent all the grades from pre-school to the university. The ninth grade contributed the

[9] Terman, L. M. *Genetic Studies of Genius*: Vol. I, Stanford University Press, 1925, pp. 648.

[10] Witty, Paul A. *A Study of One Hundred Gifted Children*: Univ. Kan. Bulletin of Education, 1930, 2, No. 7.

[11] Witty, Paul A. Some results secured in a psycho-educational clinic. *Op. cit.* p. 169.

largest number of cases. This was caused by the fact that two special studies in remedial reading were made in the ninth grade. In these studies, the facilities of the clinic were used.

Two sets of data were prepared regarding the grade placement of the cases from the first to the eighth school grades: (1) grade placement in relation to the chronological age and (2) grade placement in relation to the mental age. For these purposes Strayer's Standard of Gradation [12] was used. Since the standard used by Strayer was calculated for December, the ages of the children in this study were computed for December.

The tabulations show that 80.2 per cent of the children were of chronological age normal for their grade placement; only 9.2 per cent were underage. Witty reported 33.5, and 35.7 per cent respectively for 751 clinic cases in Kansas.

Forty-one and one-tenth per cent are of normal mental age for their grade placement, 43.9 per cent are above average, and only 14.9 per cent are underage mentally. These data, too, are strikingly unlike those set forth previously by Witty (26.3 per cent normal, 28 per cent overage, 45.7 per cent underage). More than two-fifths of the cases referred to the Northwestern clinic have mental ages which would permit them to do work at least one grade in advance of their grade placement. Two-fifths were normal in mental age for their grade placement, and less than one-fifth were underage mentally.

More than one-third (35.7 per cent) of the cases to whom reading tests were administered received scores which showed them to be one or more grades advanced for their grade placement. However, the same numbers were doing work inferior to the grade placement and 28.5 per cent were doing average work in reading. These figures stand in striking contrast to those reported for the Kansas clinic: 26.3 per cent were advanced, 57 per cent were inferior and 18 per cent were doing average work. Thus, one sees that more than half of the Kansas cases and only 38 per cent of the Northwestern cases were educationally retarded.

The location of a clinic, the community demand, and the school

---

[12] Strayer, George. *Age and Grade Census of Schools and Colleges.* U. S. Bureau of Education Bulletin, No. 5. 1911. p. 12.

practice are very important items in affecting the character of the problems referred to an educational clinic. Low mental ability played a most important rôle in causing children to be sent to the Kansas Clinic. This factor, although important in the cases sent to the Northwestern Clinic, appears secondary to the following items: reading retardation, inability to adjust in certain classes in the secondary school (wherein superior students were enrolled), inadequate study habits, and lack of interest accompanied by indifference or antagonism to the work of the school.

(3) *Emotional stability*

The case reports were used to study the amount of emotional stability displayed by the children. For purposes of comparison the two classifications that were employed by Witty are here used. "Slight tendency" toward instability includes notations such as: "predisposition," "some evidences," "a few indications"; while "marked tendency," includes: "very unstable," "marked instability," and so forth.[13]

In two hundred case studies, chosen because of the completeness of the reports and the ability of the examiners to make adequate diagnoses, a "slight tendency" was reported eighteen times (9 per cent), and "marked tendencies" appeared eighteen times (9 per cent). Combining these, we have eighteen per cent, which corresponds roughly to that reported by Witty (16.6 per cent). The National Committee upon Mental Hygiene estimates that 3.5 per cent of the school population is psychopathic. It appears therefore that the psycho-educational clinic tests an unusually large number of children who have psychopathic tendencies. Schools find that these cases present serious problems and therefore refer them to the clinic.

(4) *Remedial work*

During the existence of the Northwestern Clinic several accounts have been written. One of these which has recently been published is briefly summarized here. A group of twenty-one high school pupils entering the ninth grade at Evanston were

[13] Witty, Paul A. *Ibid. p.* 167.

placed in a special class.[14] They had been chosen from a class of 250 because of the disparity which existed between their reading ages and the averages of their educational ages and their mental ages. They were brought to the clinic and were given a series of tests which included the Binet, Porteus, Stanford Achievement Test, Form V, Gates Silent Reading Tests, Gray's Standardized Oral Reading Paragraphs, Pintner-Paterson Performance Scales, and informal tests of rate and comprehension in reading. A remedial program of instruction was planned for thirteen weeks using five forty-minute periods each week. The article fully describes the method of procedure. At the end of the thirteen weeks different forms of the educational tests were administered. The comparison of results indicated a gain of from 1 to 2.9 grade years upon each of the four Gates Tests. Considering the briefness of the time in relation to the size of the gains, one must be impressed with the possibility of securing large educational products if one has the opportunity to base his teaching upon rather thorough psychological study of each child.

PART III
CASE STUDIES

The following case studies are illustrative of the work of the clinic: Through the first, one may see the nature of the diagnostic methods followed in the clinic, and through the second, the character of a therapeutic program.

CASE A

Name: R
Grade: Sixth
Chronological Age: 11 years 9 months
Mental Age (Binet): 15 years 7 months (Med. M.A. Performance Scales 15 years)
Intelligent Quotient (Binet): 132

Achievement tests:	Subject
Reading	
Stanford—paragraph meaning	8.4
Stanford—word meaning	6.2
Gray Oral	5.1
Spelling	
Stanford—dictation	4.0
Gates—four type	5.6

[14] Witty, Paul A. and LaBrant, Lou L. Some results of remedial instruction in reading. *Educational Trends* 2:7-13, January, 1933.

R is an unusual child, whose mental alertness and general ability place him in the group designated as "gifted." He should do satisfactory work in any school situation, but it is apparent that he is now failing to utilize his ability effectively.

On the Stanford Revision of the Binet test, R achieved a mental age of fifteen years and seven months; an intelligence quotient of 132 was obtained. His basal year was twelve, although he passed all except one of the tests on the fourteen year level. Of the tests in the adult level, four were passed. R has the vocabulary of a fourteen year old child, but he is not inclined to use it either in silent reading or in oral and written expression.

R has developed an attitude of dislike and a feeling of incompetence for reading and for spelling. In the case of reading, the average results of several tests indicate that he is doing work equal to the norm for sixth grade children, although his attainment is very uneven. The summary at the top of this report reflects his superiority in comprehending the general thought of paragraphs. However, his ability in other types of reading is poor.

As measured by the New Stanford Achievement Test Dictation Exercises (spelling) R is doing work equal to that which characterizes the average fourth grade child. His inadequacy in spelling is accompanied by a distinct inferiority in oral reading and by very low attainment upon one of the Gates tests (the ability to note details). Upon the latter test R reached the norm for children of grade four. Certain of this boy's reading lacks appear to be associated with spelling deficiencies.

Upon three word lists, R's average spelling attainment was that of a fourth grade child. His compositions reflect his attainment in spelling, and they are regrettably poor in quality.

Possibility of effective remedial endeavor is diminished by R's attitude. He professes an extreme distaste for spelling and he displays an acute awareness of his particular incapacity. Because success in spelling is determined to a marked degree by attitude, this question arises: How can one change R's tendency to accept inferior products as satisfactory, and arouse in him a zest for success and an interest in producing superior products? The following procedure suggests a way: It appears desirable to

identify specific types of errors in spelling, and thereafter to provide goals which are attainable, and situations in which success will be sensed and rewarded.

In order to provide a logical start in this therapeutic work, R's spelling errors in the Ayres' lists were categoried. Errors such as the spelling in "circumstence," "essue," "desirible," illustrate clearly that *vowel difficulties and phonetic inadequacies* must be treated. In addition, R does not *differentiate syllables* (e.g. flattering was spelled flattring). He does not appear to understand the *ie* and the *ei* rule. Many of R's mistakes in written spelling seem to be caused by lapses (errors made in known words). The four categories just mentioned include many of his difficulties (80 per cent in one list).

It is suggested therefore that spelling work in the sixth grade be discontinued and that individual tutorial work be initiated. That this can prove beneficial is indicated by the results of one period (30 minutes) spent with this boy. R was asked to spell twenty words; ten of these contained the *ie* or *ei* combination. Six of the ten were misspelled. He was then requested to spell the ten aloud. Upon doing this, he quickly perceived his inconsistencies in spelling these words; he was led inductively to formulate the principle involved and to state it in his own terms. His unusual success in this hitherto profitless type of endeavor led to a change in attitude, and he became eager to continue the work. It appears therefore that a corrective program should give immediate opportunity for R to modify incorrect habits of the types listed above, and that provision should be made for strengthening underlearned habits.

R acknowledges an enthusiasm for certain aspects of history; his composition work should be articulated around this interest, and his spelling should be checked meticulously in this work. Additional integration of his interests and efforts should take place. Since interest is nourished by success, it appears necessary that the school provide occasions for R to demonstrate his rather unusual knowledge in American history.

One part of the remedial work should consist in correcting the spelling "lapses." Dictation exercises in the form of sentences should be read. Each sentence should be checked for errors im-

mediately after it is written. After accuracy, involving auto-criticism of small units, has been established, whole paragraphs should be written, and spelling should be checked at once. The habit of giving careful, analytic attention to sound-units when writing thus can be developed.

An opportunity for general improvement in spelling should be provided by assigning study lists chosen from carefully graded word lists. However, the teacher must always keep in mind the importance of sustaining interest and of providing for awareness of success.

The remedial work in reading may be done at home if necessary, and may include some such work as this: Paragraphs of materials of everyday interest, such as are found in books dealing with the social studies, in the newspapers, and in periodicals, should be read silently by R; then complete comprehension should be tested since he is especially weak in noting details. This testing may take several forms: written reactions, oral report, underlining the sentence which tells best what the paragraph is about, selecting the topic sentence, and so forth. To develop this ability further, R should be provided with stories which contain several clearly-defined episodes, each of which can be tested thoroughly. If these are related, he can be led to appreciate the fact that comprehension of later units depends upon thorough mastery of earlier ones. By this means, interest may be aroused and sustained; progress can be facilitated by the content itself. In addition, questions not precisely answered in the passage may be used. In this way attention may be directed to relevant detail. Especially appropriate would be readings which require deductions from the material given in: history, geography, science, invention, and health.

Although there is a decided unevenness in R's reading attainment, several limitations stand out clearly. R needs much remedial work in the following types of reading: understanding precise directions, reading intensively, getting every factor accurately, and relating items to the general theme.

The examiner believes that the chief basis of motivation resides in providing for this boy abundant reading materials *which relate to his special interests and which are in accord with his ability.* Of course, attention should be directed to particular errors in

spelling and in reading, and a graded, orderly program of remedial endeavor must be instigated. Nevertheless, most important of all is the problem of attitude which can be, it appears, radically changed by the methods herein suggested.

CASE B

This case illustrates the complexity of the reading function as well as the fact that remedial work must provide motivated drill in which goals are set up which are in accord with the child's ability and interest.

J was in 7A; his IQ was 93. His responses to the Stanford-Binet test were uniformly slow, labored, and inadequate. An amiable disposition and a willingness to attempt *any* task made his deficiency less noticeable at first. His vocabulary rating was equal to the norm for 10 year old children in the oral vocabulary section of the intelligence test. J's responses upon the Porteus test were noticeably slow. His IQ calculated from the Porteus Maze Tests was 90. He received a rating of "normal" in the physical tests, and no handicap was found either in his vision or in his hearing. He was right-handed.

Upon the educational tests several features of his performances were conspicuous. His reading age was that of a fifth grade boy upon the Stanford Achievement Test, Paragraph Meaning, but his vocabulary score was equal only to that of an average fourth grade child. His composite score upon the Gates tests corresponded to the norm for 5.3 grade; however, poorest attainment was exhibited in "ability to note details." His responses upon the Gray Oral Reading Paragraphs revealed a weak vocabulary; and mispronunciations, substitutions, and omissions were frequent. J reached the fourth grade norm in oral reading and his rate of silent reading was 120 words per minute. He appeared not at all disturbed by the fact that many words were skipped over both in silent and in oral reading, and held no meaning whatever for him. He exhibited an indifference for reading, but he was unlike most poor readers in that he did not display an antipathy toward the task. However, he had not read voluntarily a single book during the six months preceding the examination.

J's difficulties were numerous. They included: meager back-

ground of experience; inadequate vocabulary; inaccurate recognition; inability to deal successfully with large reading units; phonetic lacks; and indifference.

He was interrogated concerning his interests and his vocational ambition. He was actively interested in aviation, and expressed an ambition to be an aviator. Surprising indeed was the information (far exceeding that of the examiner) which the boy possessed and eagerly displayed concerning motors and airplanes.

A graduate student worked daily with J for eighteen weeks. Phonetic drills and vocabulary lessons (on fifth grade level) were planned. These were accompanied by a provision for supplementary reading (at first about aviation). The supplementary reading was checked carefully for details, and J was encouraged to note and react to sensory images as he read. Drill with especially devised flash cards and with phonetic materials called attention to the essential characteristics of new words. After two weeks of this type of drill, a fifth grade reading book was introduced, and the large unit was made basic during the reading. A fifth grade reader was now used for rapid reading drills. Five weeks later, a book appropriate in the sixth grade was introduced and was used in rapid reading exercises. Every reading unit was checked thoroughly for comprehension.

Three important supplements to the reading work deserve emphasis. The new words introduced during the reading drills were used in J's spelling lessons; the social science teacher devoted time to J's special interest (air-travel), allowing him to lead and direct the discussions; and in the English work he was permitted to write about his favorite subject. Thus, most of the school work was organized about a central, interesting theme; the boy had the joy of "sensed" success; and adaptation of all materials was effected in terms of the boy's maturation levels.

After eighteen weeks, J was reading effectively some seventh grade materials and many sixth grade books. In addition he was doing abundant voluntary reading. A gain of 1½ years was registered in the composite score upon the Gates tests, and two years' improvement was exhibited upon standardized spelling scales. Noticeable but quantitatively unmeasured was the improvement in composition writing. Clearly observed also was

the therapeutic value of the program upon the mental hygiene of this boy. J had been transformed from an indifferent drone into an active, interested participant in school work.

THE EFFECT OF COLLEGE FRATERNITIES ON SCHOLARSHIP

T. M. CARTER

Albion College

THE men studied were divided into two distinct sections, fraternity and non-fraternity. In order to be listed as a fraternity man, the student must have spent his entire four years in Albion College, and must have been a member of a fraternity during the entire time. In order to be listed as a non-fraternity man the student must have spent his four years at Albion College, and must not have been a member of a fraternity during any of that time. This criterion of selection was so rigid that only 114 fraternity and 65 non-fraternity men were available from the last seven graduating classes. We could go no farther back than this for some of the data needed in our study were not available for classes earlier than these.

The general method was to arrive at some index of promise based upon the student's high school record and the score on the intelligence test taken upon entrance to college, and then to find his index of achievement while in college. The relationship between these two indices would, we assumed, show whether life in a fraternity or out of it was most conducive to scholarship.

The high school record was evaluated by assigning to the grades A, B, C, D, and E, numerical values from one to twelve as follows:

A	12
A−	11
B+	10
B	9
B−	8
C+	7
C	6
C−	5

D+	4
D	3
D−	2
E	1

The average of these grades gave the student's high school index.

In securing the intelligence test index we divided the entire spread of mental test scores for each freshman group into twelve equal steps or intervals comparable to the twelve high school achievement intervals. The interval in which a pupil's score fell determined his mental test index. The average of the student's high school index and his mental test index, we called his index of promise.

In arriving at an index of college achievement the student's A, B, C, D, and E grades were given the same numerical values as they were given in determining high school index.

The first use made of our data was to find the correlation between the index of promise and the index of achievement for the fraternity men and the correlation for the same variables of the non-fraternity men. The coefficients of correlation obtained are as follows:

For the fraternity men	.68
For non-fraternity men	.59

However, upon brief inspection, it was obvious that the coefficient of correlation did not show very adequately the thing we wished to find, namely, whether life in a fraternity was beneficial or detrimental to scholarship. The higher coefficient of correlation for the fraternity men shows only that the scholarship of the fraternity men conformed more closely to the index of promise than did the scholarship of the non-fraternity men. Non-conformity by making an index of achievement higher than the index of promise would tend to lessen the coefficient of correlation just as much as non-conformity by making an index of achievement less than the index of promise.

We next used the method of averages and it was found that the average index of promise for the fraternity men was 7.80 and

that the average index of promise for the non-fraternity men was 7.97. The difference here is small but in favor of the non-fraternity men.

The average index of achievement for fraternity men was found to be 8.19 and that of the non-fraternity men 8.42. Here again the advantage is in favor of the non-fraternity men but is not large enough to be significant. Thus it will be seen that the students who joined fraternities did not have indices of promise significantly different from the indices of promise of those students who did not join fraternities and that the averages of the indices of achievement of the two groups are very nearly the same.

The variation of the students' achievement from their own respective indices of promise was next investigated. We found that the average variation of the indices of achievement from the indices of promise of the fraternity men was 1.31 and that of the non-fraternity men was 1.48. For the fraternity men the average of the plus variations was 1.52 and the average of the minus variations was 1.04. For the non-fraternity men the average of the plus variations was 1.62 and the average of the minus variations was 1.24. Thus it appears that the variations of the indices of achievement of the non-fraternity men from their indices of promise are greater than are the same type of variations of the fraternity men. This only shows, from another angle, what was shown by the above coefficients of correlation, namely, that there is a closer correspondence of the index of achievement with the index of promise for the fraternity men than for the non-fraternity men.

We next tried to determine whether the effects of fraternity life were greater on the students with high indices of promise or on those with low indices of promise. The following brief tables show our findings in this regard:

TABLE I

Achievement of Students Having Indices of Promise Above the Average

	Fraternity	Non-fraternity
Achievement greater than promise	28%	50%
Achievement less than promise	72%	50%

TABLE II

Achievement of Students Having Indices of Promise Below the Average

	Fraternity	Non-fraternity
Achievement greater than promise	84%	77%
Achievement less than promise	16%	23%

The tables show that of the students with indices of promise above the average the non-fraternity men make a better showing than the fraternity men. Fifty per cent of the non-fraternity men with high indices of promise make indices of achievement higher than their indices of promise as compared to twenty-eight per cent of the fraternity men who can show such results.

Of the students having indices of promise less than the average, those belonging to fraternities make a better showing. Eighty-four per cent of this group have indices of achievement higher than their indices of promise as compared to seventy-seven per cent of the non-fraternity men.

In attempting to determine the cause of the difference in achievement between fraternity and non-fraternity men, we secured the cooperation of 79 men (39 fraternity and 40 non-fraternity) to the extent that they kept a record of how they spent their time for a period of one week. The instrument used by the student in keeping this record was so devised that every moment in each 24 hour period was accounted for. A complete summary of these findings has been published as a part of a more comprehensive report of how college students spend their time;[1] thus only the findings relating to comparisons between fraternity and non-fraternity men will be referred to here. The data show that the non-fraternity men spend more time in curricular activities and in remunerative employment. The fraternity men spend more time in extra curricular activities, in reading for pleasure, and in personal activities. By personal activities was designated a number of things such as personal correspondence, dressing, bathing, barbering, caring for clothing, personal shopping, nap-

[1] Carter, T. M., "Activities of College Students," The PHI DELTA KAPPAN, August, 1932.

ping, day-dreaming and doing nothing so far as purposeful activity is concerned. The group of students who participated most in extra curricular activities was composed of fraternity men whose indices of promise were above the average. This, in my judgment, helps to explain why this particular group, in proportion to index of promise, did less well than any other group.

It would be desirable to know to what extent fraternities hold men in college. In our attempt to answer this question we found that the median of the annual withdrawals over a period of years for the fraternity men was 19.5 per cent and for the non-fraternity men was 26.5 per cent. If we assume all other factors to be equal, the fraternities could then be credited with a noticeable degree of holding power. However, it is quite likely that not all of the other factors are equal. Whether or not other factors are unequal to the extent of accounting for all the differences here represented, is a matter upon which further data are needed.

These comparisons lead to a very interesting observation not directly related to our original undertaking but which may have some educational significance. Of the students having indices of promise above the average, thirty-three per cent have indices of achievement greater than their indices of promise. While of the students who have indices of promise less than the average, eighty-

TABLE III

Achievement Quotients for Ten Students

Pupil No.	Index of Promise	Index of Achievement	Achievement Quotient
1	5.30	6.57	124
2	3.75	6.10	163
3	5.20	6.90	133
4	5.05	6.97	138
5	4.80	6.40	133
6	10.60	7.82	74
7	9.90	7.87	79
8	9.75	7.43	76
9	9.95	7.84	80
10	9.40	7.83	83

one per cent have indices of achievement greater than their indices of promise.

Some of the variations of indices of achievement from indices of promise are quite large. Samples of the more striking ones are seen in Table III.

In terms of absolute achievement the first group of five students is not doing so well as the second group, but each student has an index of achievement considerably higher than his index of promise and each has an achievement quotient considerably above one hundred.

The second group of five students is making a better absolute showing but relatively their work is inferior to that of the first group. The achievement quotient of each student in the second group is considerably less than one hundred.

What can these facts mean? Any true interpretation leads us at once into the problem of the much debated meaning of A.Q. The technique used in this study makes possible the obtaining of an A.Q. at the college level. This A.Q. is obtained by dividing the index of achievement by the index of promise.

The A.Q. has been the subject of a large number of articles which have appeared in various journals. Possibly the most thorough analysis and evaluation of this quotient has been made by Toops and Symonds.[2] These authors have shown that "the implications of A.Q. have been subject to a great deal of confusion even among the originators of the technique." Perhaps our derivation of A.Q. and the implications deduced from our ratio will only add to the confusion. However, one thing seems generally agreed upon, namely, that the purpose of the A.Q. is to show the degree to which a pupil's actual progress has attained to his potential progress. This has been done in the past by finding the relation between the mental age and educational age. Mental age has been used to represent potential progress while educational age was allowed to represent actual progress. In our case potential accomplishment is determined by two factors, namely, high school grades and the relative value of the score on mental

---

[2] Toops, H. A., and Symonds, P. M. "What Shall We Expect of the A.Q.?" Journal of Educational Psychology, Vol. XIII, p. 513.

tests. Actual accomplishment is determined by grades made while in college. Can our A.Q. be taken to represent a measure of effort or motivation? If so, the students with high indices of promise, generally speaking, are not being motivated nearly so much as students with low indices of promise. Freeman[3] contends that the A.Q. does not warrant making such interpretations. He suggests that "when we classify individuals according to their IQ and expect them to achieve in proportion we do a certain injustice to the individual of high intelligence." He says that "pupils of high IQ must have lower A.Q.'s as a consequence of the fact that the ability required in the intelligence test and the ability required for achievement in school will not be identical." This criticism is doubtless justified; however, it does not apply with as much force to our A.Q. as it does to those A.Q.'s which are differently derived, since the denominator of our quotient is determined only in part by the intelligence score and in part by past achievement.

It has been reported that "Dr. May,[4] as a result of an investigation of the relationship between hours spent in study, intelligence and school marks of college students found that in case of his students, there is a decided negative correlation between hours spent in study and intelligence. This can only mean that many bright students able to 'get by the passing mark' with little study, prefer to spend a proportionately large part of time on other than study activities." Our study shows the same trend, and also shows that this tendency is greater among students who join fraternities than among students who do not join fraternities.

SUMMARY AND CONCLUSION

1. The average index of promise of students who join fraternities is not significantly different from those who do not join fraternities.

2. The average achievement of fraternity men is almost exactly the same as the average achievement of the non-fraternity men.

3. Students having indices of promise higher than the average

---

[3] Freeman, F. M., Mental Tests, Houghton, Mifflin Co., 1926.

[4] May, M. A., Journal of Educational Psychology, Vol. XIII, p. 528.

are more likely to make indices of achievement higher than their indices of promise if they do not join fraternities.

4. Students having indices of promise less than the average are more likely to attain indices of achievement higher than their indices of promise if they join fraternities.

5. Generally, students, whether fraternity or non-fraternity, who have low indices of promise are more likely to make indices of achievement higher than their indices of promise than are the students who have high indices of promise.

TRAINING FOR VOCATIONS IN THOSE OF HIGH AND LOW INTELLIGENCE

LOUIS F. HILLMAN

Indiana University

INTELLIGENCE quotients of the Indiana University graduates are not available but intelligence ratings are. The intelligence ratings are available for only a part of the 1928 graduating class. The *Thurstone Psychology Examination, Test IV*[1], was given to the entering freshmen for the first time in the fall of 1923. It was optional; consequently only a few of that year's freshmen took the examination. The first graduating class that could have such intelligence ratings on record in the Psychology Department is the graduating class of 1928. The Department of Psychology gave the examination, did the scoring, and reduced the raw score made by those who took it to a percentile rating.

These percentile ratings of the raw score of the *Thurstone Psychology Examination, Test IV,* are used in this study instead of intelligence quotients.

In order to use the data to answer the question, those of the graduating class of 1928 who have percentile ratings on the Thurstone test have to be grouped according to their percentile ratings into four equal groups. In order to do this the percentile ratings are assembled into a frequency distribution table. Then the 25th, 50th, and 75th percentile points are found. Table I shows the number, the percentile ratings of the raw score, the range, and the 25th, 50th, and 75th percentile points of those of the 1928 graduating class who took the *Thurstone Psychology Examination,* Test IV, as freshmen.

---

[1] The Thurstone Psychology Examination, Test IV, for College Freshmen and High School Seniors.

TABLE I [2]

*Frequency Distribution Table of the Number, Percentile Ratings of the Raw Score, the Range of These Ratings, and the 25th, 50th, and 75th Percentile Points of the 1928 Graduating Class Who Took the Thurstone Psychology Examination, Test IV , as Freshmen**

Percentiles	Frequencies	Quartile points
100		
90	17	
80	14	
70	28	Third quartile = 77.9
60	15	Second quartile = 60.3
50	17	
40	16	First quartile = 37.9
30	16	
20	15	
Less than 20	9	
Total	147	

* Table reads: Seventeen of the graduates have percentile ratings between 90 and 100. The 75th percentile point is 77.9, etc.

Table I shows that 147 of the members of the 1928 graduating class took the *Thurstone Psychology Examination,* Test IV, as freshmen. The lowest percentile rating made by any of those who took the examination is eight, and the highest is 98. Of the 147 graduates, 17 have a percentile rating from 90 to 100; and nine have a percentile rating of less than 20. The mode is the step interval between 70 and 80. The median, or 50th percentile point, is 60.3; the 25th percentile, or first quartile point, is 37.9; and the 75th percentile, or the third quartile point, is 77.9. About one-fourth of the 147 graduates should have percentile ratings falling between 77.9 and 100; one-fourth should have percentile ratings falling between 60.3 and 77.9; and the other half of the 147 graduates should be evenly divided between points of 37.9 and 60.3, and below 37.9.

[2] Hillman, Louis F. Do Indiana University Graduates Follow A Vocation For Which They Are Trained? page 299 (an unpublished doctor's thesis on file in the library of Indiana University).

Table II shows the distribution of the 147 Indiana University graduates of the 1928 graduating class who took the *Thurstone Psychology Examination,* Test IV, as freshmen, into four groups according to the percentile ratings.

Thirty-nine of this group made raw scores on the test which made their percentile ratings between 100 and 77.9; and 37 made percentile ratings between 77.9 and 60.3; thirty-seven of the graduates have percentile ratings between 60.3 and 37.9; and 34 of them fall in the lowest group of intelligence.

TABLE II [3]

The 147 Indiana University Graduates of the Class of 1928 Who Took the Thurstone Psychology Examination, Test IV, Distributed into Four Groups According to Percentile Ratings *

Percentile ratings	100- Number	77.9 Per cent	77.9- Number	60.3- Per cent	60.3- Number	37.9 Per cent	37.9- Number	0 Per cent
Follow vocation for which trained	24	61.5	22	59.4	26	70.3	24	70.5
Do not follow vocation for which trained	15	38.5	15	40.6	11	29.7	10	29.5
Total	39	100	37	100	37	100	34	100

* Table reads: Out of the 39 who fall in the highest intelligence group, 24, or 61.5 per cent of the group, follow a vocation for which their Indiana University training is training; of the 37 who fall in the second to the highest intelligence group, 22, or 59.4 per cent, follow a vocation for which they are trained at Indiana University; etc.

A summary of the above follows:

4Q group: 100-77.9, has 39 of the 147 graduates.
3Q group: 77.9-60.3, has 37 of the 147 graduates.
2Q group: 60.3-37.9, has 37 of the 147 graduates.
1Q group: 37.9- 0, has 34 of the 147 graduates.

Q stands for quarter.

[3] Hillman, Louis F. Do Indiana University Graduates Follow A Vocation For Which They Are Trained? page 300.

The table shows that of the 39 who rank highest in intelligence, according to the *Thurstone Psychology Examination,* Test IV, 24, or 61.5 per cent, are following a vocation for which the training at Indiana University is training, but the table also shows that 15, or 38.5 per cent, of the highest intelligence group do not use the training in their daily vocations. Of the 37 whose intelligence rating puts them in the group second to the highest in intelligence, 22, or 59.4 per cent, use their Indiana University training in the pursuit of their life vocations. The group that ranks second to the lowest in intelligence does even better than the group which ranks highest in intelligence, for 26 of the group, or 70.3 per cent of its members, are engaged in a vocation for which the Indiana University training is training. The group that ranks lowest in intelligence has more of its members engaged in vocations for which the Indiana University training prepared them than any of the other three intelligence groups. The four intelligence groups are ranked below with respect to the per cent of their members who follow vocations for which they are trained:

1Q group: 70.5 per cent.
2Q group: 70.3 per cent.
4Q group: 61.5 per cent.
3Q group: 59.4 per cent.

It should be remembered that the *Thurstone Psychology Examination,* Test IV, is sometimes given to high school seniors. If college freshmen have had experience with the examination before they enter the University, they are apt to score higher than if they have had no such previous experience.

According to the *Thurstone Psychology Examination,* Test IV, which is a group intelligence test, the graduates with the highest intelligence do not use their training in their daily vocations as much as do those with low intelligence ratings. To those with intelligence ratings below the median the Indiana University training has more vocational value than to those whose intelligence ratings put them above the median of intelligence. It may mean that people with low intelligence have to depend more on their training to make a living than do those with the higher intelligence. According to the data obtained through the *Thurstone*

Psychology Examination, Test IV, the higher the intelligence of the individual the less he has to depend on his training to make a living.

MEN

The case is slightly different with the two sexes. Table III presents the case with reference to the men.

TABLE III [4]

The Men of the 147 Indiana University Graduates of the Class of 1928 Who Took the Thurstone Psychology Examination, Test IV, Distributed into Four Groups According to Percentile Ratings *

Quartile points	100- Number	77.9 Per cent	77.9- Number	60.3 Per cent	60.3- Number	37.9 Per cent	37.9- Number	0.0 Per cent
Follow vocation for which trained	12	80	12	75	17	81	11	61.1
Do not follow vocation for which trained	3	20	4	25	4	19	7	38.9
Total	15	100	16	100	21	100	18	100

* Table reads: Out of the 15 men in the highest intelligence group, 12, or 80 per cent, follow a vocation for which the Indiana University course is training, and three, or 20 per cent of that group, do not use the training; etc.

It should be noted that a majority of the men fall below the median of intelligence of the 147 men and women, and that the women of the group have a majority above the median of intelligence.

Table III reveals that out of the 147 Indiana University graduates of the class of 1928 who took the *Thurstone Psychology Examination,* Test IV, 70 are men and 77 are women. When the 70 men are distributed into four groups according to their percentile rating, they are as follows:

[4] Hillman, Louis F. Do Indiana University Graduates Follow A Vocation For Which They Are Trained? page 303.

4Q group: 15.
3Q group: 16.
2Q group: 21.
1Q group: 18.

It will be observed from Table III that 12 of the 15, or 80 per cent of the 4Q group in intelligence, are following their majors as a preparation in the vocation they are following. The group that ranks second to the lowest in intelligence has 17 of its 21 members, or _ per cent, who follow a vocation for which they are trained. The group that ranks as the lowest fourth in intelligence has 11 of its 18 members, or 61.1 per cent, follow vocations for which they are trained while at the University. When these four intelligence groups are ranked with respect to the per cent who follow vocations for which they are trained, they fall as follows:

2Q group: 81 per cent.
4Q group: 80 per cent.
3Q group: 75 per cent.
1Q group: 61.1 per cent.

What seems to be true of the group as a whole is not true of the men. Of the men, the number of the group that ranks lowest in intelligence, according to the *Thurstone Psychology Examination,* Test IV, use their Indiana University training least in their daily vocation. The group that ranks just below the median in intelligence has a larger per cent of its members employed in a vocation for which they are trained than any of the other groups. It cannot be said of the men that the higher the intelligence, the less vocational value has the Indiana University training.

WOMEN

The situation with reference to the women is presented in Table IV.

TABLE IV [5]

The Women of the 147 Indiana University Graduates of the Class of 1928 Who Took the Thurstone Psychology Examination, Test IV, Distributed into Four Groups According to Percentile Ratings *

Quartile points	100-77.9 Number	Per cent	77.9-60.3 Number	Per cent	60.3-37.9 Number	Per cent	37.9-0.0 Number	Per cent
Follow vocation for which trained	12	50	10	47.6	9	56.3	13	81.2
Do not follow vocation for which trained	12	50	11	52.4	7	43.7	3	18.8
Total	24	100	21	100	16	100	16	100

* Table reads: Of the 24 women in the highest intelligence group, 12 follow a vocation for which the training at Indiana University is training, and 12, or 50 per cent do not use their training in that way; etc.

Table IV reveals that 77 of the 147 Indiana University graduates of the class of 1928 who took the *Thurstone Psychology Examination,* Test IV, are women. When the 77 women are distributed into four groups according to their percentile ratings, there are the following in each of the intelligence groups:

4Q group: 24.
3Q group: 21.
2Q group: 16.
1Q group: 16.

It will be observed that of the 24 women in the highest intelligence group, 12, or 50 per cent, follow a vocation for which their Indiana University training has trained them. Of the women in the second to the highest intelligence group, 10, or 47.6 per cent, are using their training in their work. Of the women in the second to the lowest in intelligence, nine, or 56.3 per cent, are engaged so that their training is needed. The group of women

[5] Hillman, Louis F. Do Indiana University Graduates Follow A Vocation For Which They Are Trained? page 305.

that stand lowest in intelligence has 13, or 81.2 per cent, of its members using their training in the pursuit of their daily work. Ranking the four groups with respect to the per cent of the members of the group that follow vocations for which they are trained, they fall in the following order:

1Q group: 81.2 per cent
2Q group: 56.3 per cent
4Q group: 50 per cent
3Q group: 47.6 per cent

According to the intelligence ratings on the *Thurstone Psychology Examination,* Test IV, the more intelligent the women, the less do they use their training in their daily occupations.

SUMMARY

From the data derived from the *Thurstone Psychology Examition,* Test IV, it appears:

1. The higher the intelligence of an individual, the less likely is he to follow a vocation for which he is trained.

2. This is more pronounced in the case of the women, where the majority are above the median of intelligence, than in the case of the men where the majority are below the median of intelligence.

AN EXPERIMENTAL STUDY OF PERSISTENCE

ROBERTA CRUTCHER

University of Southern California

THE purpose of this study was to investigate the nature of persistence or the tendency to continue certain definite activities in the face of difficulties, and to determine the possibility of its measurement. An attempt was made to use the experimental method with control by statistical devices to see if exact results could be obtained in the study of this personality trait. In order to study persistence objectively certain difficult tasks were set the subjects and the time they continued to work at these tasks was taken as evidence of their persistence.

Six tasks were chosen with the intent to minimize as far as possible the effect of specific interests:

1. *Card Houses* (manual dexterity). The subject was presented with an ordinary deck of playing cards and asked to build a house of two stories.

2. *Bolt and Nut Puzzle* (mechanical skill). This is a "trick" puzzle consisting of a screw and a nut with a ring between the head of the screw and the nut, and the task is to get the ring off the screw.

3. *The Three Nags Puzzle* (mechanical skill). This is an ordinary mechanical puzzle, consisting of a small metal bar with a ring at each end and three loops with nags' heads on them, the task being to remove the center loop from the bar.

4. *Addition Sums* (facility with numbers). This task was to add two place columns of four figures.

5. *Copying a Picture* (artistic ability). The subject was presented with a complicated picture and asked to draw it.

6. *Cancellation of A's* (routine activity). Ordinary printed material was presented to the subject and he was asked to cross out all the a's he could find.

The six tasks were presented to all the subjects individually in

the order given. The 83 subjects were four groups of children from London Schools: (1) for physical defectives; (2) for mental defectives; (3) for orphaned children; (4) for normal children.

TABLE I

Table showing the distribution of scores for each group and each test.

		Schools				
Tests	Minutes	Feeble-minded	Normal	Orphan	Physical Defective	Total
1. Cards	0- 3.9	5	1	10	4*	20
	4- 7.9	8	9	6	9	32
	8-11.9	5	5	2	2	14
	12-15.9	0	5	2	1	8
	16-	2	2	2	2	8
2. Bolt	0- 3.9	12†	10	11	11†	44
	4- 7.9	5	9	7	5	26
	8-11.9	1	1	3	0	5
	12-15.9	1	2	0	0	3
	16-	0	0	1	1	2
3.3 Nags	0- 3.9	12	6	12	8	38
	4- 7.9	2	5	6	7	20
	8-11.9	3	7	0	4	14
	12-15.9	3	1	1	0	5
	16-	0	3	3	0	6
4. Add.	0- 3.9	4	0	1	2	7
	4- 7.9	5	5	9	6	25
	8-11.9	4	9	2	7	22
	12-15.9	2	2	4	2	10
	16-	5	6	6	2	19
5. Picture	0- 3.9	7	1	2	2	12
	4- 7.9	6	9	9	7	31
	8-11.9	4	7	6	6	23
	12-15.9	0	3	3	0	6
	16-	3	2	2	4	11
6. Cancellation	0- 3.9	4	0	5	2	11
	4- 7.9	6	8	7	10	31
	8-11.9	6	9	5	4	24
	12-15.9	1	3	3	1	8
	16-	3	2	2	2	9

*Task refused because of paralyzed hand.

†Task omitted because puzzle not working right.

Each group was about evenly divided between boys and girls, the ages ranging from about seven years to sixteen years. The mental ages were seven to thirteen. The majority were eight, nine, and ten in mental age. The time the child persisted at each task was taken with a stop watch and recorded, with a time limit of twenty minutes for each task.

Results from the four different groups are shown in Table I. The total scores for each group show that the greatest number of cases falls in the second lowest group of time scores for each school except the feebleminded where it falls in the lowest. Aside from this there are no noteworthy differences between the groups. There is a more noticeable difference between the tests. The scores for tests two and three (mechanical puzzles) are much lower than for the other tests. The results taken as a whole show that mental or physical defect alone does not influence to any marked extent the reactions to this experiment.

TABLE II

	Average Total Scores					Average IQ	
	Boys	No.	Girls	No.	Boys / %Girls	Boys	Girls
N.	3972″	10	2691″	12	148	118	111
R.	3121″	10	2870″	12	109	89.4	93.9
M.D.	2951″	10	2470″	10	120	63.5	63.3
P.D.	2940″	10	2670″	9	110	87.7	93

N.—Normal elementary school.
R.—Residential orphan school.
M.D.—School for mental defectives.
P.D.—School for physical defectives.

Table II gives the average total scores for the boys and girls of each school. In every case the score for the boys is higher but the differences are not very great in any but the elementary school. This difference cannot be due solely to a difference in IQ, for, as Table II shows, the greatest difference in IQ between the boys and girls is only one of seven points, where the difference in scores is 48 per cent. In the physical defective school there is

a difference of 5.3 points in IQ in favor of the girls and a difference in total scores of only 10 per cent in favor of the boys.

TABLE III

Average Scores on Each Task.

	No.	1. Cards	2. Bolt	3.3 Nags	4. Add.	5. Picture	6. Cancel
Boys	40	489″	286″	415″	767″	584″	588″
Girls	43	433″	255″	377″	522″	477″	502″

In Table III are the average scores on each task for boys and girls. Again the boys' scores are higher in every case than the girls' but in most cases very little higher. The only task that shows any marked difference for the two sexes is the Arithmetic, in which the boys persist longer by four minutes than the girls. The results as a whole do not indicate that the boys are more persistent than the girls.

The correlation of scores with intelligence quotients for the group as a whole is plus .30. The results show that the lowest scores were made by the defectives but that one defective had the sixth highest score. The children of superior intelligence are to be found in the medium and high score groups. The child with the highest IQ of the group (138) has a fairly low score (2185″). The total scores range from 430″ to 6690″. The five lowest scores (from 430″ to 955″) are those of defectives. The highest score (6690″) is that of a boy with an IQ of 103. The next highest (6060″) is that of a boy with an IQ of 129. The conclusion must be that whatever was measured by the experiment was not general intelligence alone but something which is positively related to it.

The statistical method has been applied to the results in order to learn something about the factors involved in persistence. There are six different tasks in all and in order to find if there is a general factor involved in these tasks, the results of each task have been correlated with those of every other task. This results in fifteen correlation coefficients as shown in Table IV. They are all positive correlations and some of them are fairly high. This would seem to suggest that there might be some factor common to all of the tasks. It certainly indicates that there are other factors

TABLE IV

Intercorrelations of the Six Different Tasks

	1. C.H.	2. B.N.	3. 3N.	4. A.	5. C.P.	6. C.A.
1.C.H.		.45	.56	.48	.49	.49
2.B.N.	.45		.71	.37	.29	.41
3.3N.	.56	.71		.47	.23	.50
4.A.	.48	.37	.47		.33	.64
5.C.P.	.49	.29	.23	.33		.30
6.C.A.	.49	.41	.50	.64	.30	

not common to but different for each task. In order to determine whether there is a general factor and specific factors or only many specific factors, the formula of the tetrad equation has been applied and the tetrad-differences ascertained. (The method, use and applications of the tetrad-difference are explained in Spearman's *The Abilities of Man,* p.73ff. and in the Appendix.)

If the mean of the tetrad-differences is approximately the same as the expected probable error, a general factor is present in the tasks. If the mean is three times as large as the expected probable error, no general factor is present in the tasks but only related specific factors or "group factors" (see Spearman's *The Abilities of Man,* chap. XIII).

Mean of 45 tetrad-differences	.0821
Observed p.e., from *mean* x .8453	.0694
Observed p.e., from sigma	.0723
Theoretical PE	.0377

Since the mean here is only about twice the size of the expected probable error, there must be factors present which are shared by some tasks and not by others and it may be possible to locate the group factors, eliminate them, and then determine if there is present a general factor.

It is supposed that the two mechanical puzzles involve a factor which is common to them but not to any of the other tasks or not to such an extent. This expectation is borne out by the comparatively very high correlation (.71) found between them. If this correlation is omitted, and also the tetrad-differences involving

it, from the calculations, it is found that the mean (.0572) more closely approximates the theoretical probable error. Still there is a discrepancy. If the other comparatively high correlation (.64) is eliminated, the values approximate one another still more closely:

Mean of 23 tetrad-differences	.0423
Observed p.e., from *mean* x .8453	.0358
Observed p.e., from sigma	.0378
Theoretical PE	.0342

These results do not prove either the presence or the absence of a general factor, but since the difference between the mean and the theoretical probable error is smaller when the correlations which most clearly seem to include group factors are eliminated, the presence of a general factor is indicated.

The conclusions must be that there are group factors in the tasks which partly, if not entirely, account for the positive correlations between them. With certain of the group factors eliminated the results begin to look as if there might be present a general factor and specific factors as well. In any case what is measured is not a simple, unitary quality but a complicated element, or a complication of elements.

In order to determine what part interest or preference played an attempt was made to elicit from each child which task he "liked best" and which, "next best." It was felt that the answers to these questions were highly unreliable. The child usually hesitated for a long time and then seemed to choose at random from the tasks which were spread out before him. Some children refused to make a choice; some made only one choice. Out of first choices made by seventy-three children, twenty-five of them coincide with the tasks of highest "persistence" score, or in other words, twenty-five children preferred the task that they worked at longest, seventeen selected the one they worked at next longest, and so on. The second choices (made by sixty-six children) were probably even less reliable, but here again the greatest number of preferences coinciding with any one score is that which coincides with the highest score, so there is some indication that the child prefers the tasks that he works at longest.

TABLE V

Comparison of preference with "persistence" score.

Rank in "Persistence" Score	1. Preference	2. Preference
1.	25	16
2.	17	9
3.	9	10
4.	6	14
5.	9	8
6.	7	9

TABLE VI

Popularity of Each Task.

Task	1. Choice	2. Choice
1.	15	14
2.	2	1
3.	10	12
4.	26	17
5.	13	10
6.	7	12

The most popular task by far was Addition (No. 4). The least popular task was the Bolt and Nut Puzzle (No. 2), taken only twice as first choice and once as second. Table VI shows the number of times each task was preferred.

From notes made on the child's behavior during the experiment certain conclusions may be drawn about the motives involved in "persistence." Stated briefly they are: *A*) interest, as evidenced by complete engrossment in a task; *B*) a certain amount of success on a task; *C*) a definite feeling of unwillingness to admit defeat or failure, shown by the invention of such excuses as a feeling of fatigue, a sore hand, etc., for giving up rather than admitting failure; *D*) a desire to please the experimenter or to gain attention and praise; *E*) perseveration or inertia.

Finally, it was noted that one type of behavior characterized the "persistent" child. There were some children who stood out as being particularly diffuse—talkative, active and restless,—and there were some who were particularly silent and quiet, answering questions briefly and hesitatingly. There were nine chosen as especially talkative and nine as especially silent. Of these latter nine, only one has a low score. The five highest scores are included in this group and they rank as follows: 18, 54, 55, 73, 81, 82, 83, 84, 85. This indicates that the highest "persistence" scores are gained by those children who show more than an average tendency to be quiet and subdued.

There are no characteristic differences in scores between the mental defectives, physical defectives, orphans and normal school

children. They tend to show a similar scatter with a grouping of the mental defectives in the low persistence scores. There is a relationship between general intelligence and the results of the test to the amount of plus .30. The only marked differences between the two sexes is the higher score for boys on the Arithmetic test. This alone does not justify us in concluding that boys are more persistent than girls. The experiment measures not one general trait but a combination of factors, probably including general intelligence, special abilities, perseveration and determination. These factors seem to be closely related to one another and are not easily separable into isolated, independent traits. Interest or preference seems to have influenced the results in a certain measure as shown by the tendency to prefer the test that was worked at longest. The most "persistent" children seemed to be those who talked the least, who were the quietest, who belong perhaps to the "introverted" type.

In this experiment only one measure of persistence has been used, total time of working at a difficult task. Other measures of persistence are to be found in previous experiments. June Downey (2) took the control of handwriting as a measure of persistence, but the results are limited by the use of only this one activity. Fernald (3) measured persistence by the length of time a person can stand on his toes, which seems to be largely a measure of physical strength. Chapman (1) experimented on the relation of persistence to speed of work and success. His results show certain factors which seem to enter into persistence, namely: (1) type of training received by the pupil; 2) prestige of the examiner; 3) setting of the test; 4) manner in which the subjects weight the importance of the result. Morgan and Hull (4) judged the persistence of each subject by his general attitude and attack on alley mazes of increasing difficulty. Their judgments showed a high correlation with ratings made by personal friends of the subjects. Webb (5) has also reported a study on "will" based on answers to questionnaires. From the data he concluded that there is a general factor present on the "character" side of mental activity as well as on the "intellective" side.

From what has already been reported comes the suggestion that in the future not only should time and errors be measured but

attempts should also be made to ascertain and measure motives and attitudes.

REFERENCES

(1) Chapman, J. C. "Persistence, Success and Speed in a Mental Test." *Pedagogical Seminary*, 1924, XXXI, p. 276.

(2) Downey, June. *Will Temperament and Its Testing*. The World Book Company, 1924.

(3) Fernald, G. G. "Character versus Intelligence in Personality Studies." *Journal of Abnormal Psychology*, 1920, XV, p. 1.

(4) Morgan, J. J. B. and Hull, H. L. "The Measurement of Persistence." *Journal of Applied Psychology*, 1926, X, p. 180.

(5) Webb, E. "Character and Intelligence." *The British Journal of Psychology, Monograph Supplement*, 1915, I, pt. 3.

A TEST OF EYE-HAND COORDINATION AT THE FIFTH YEAR LEVEL

LORRAINE B. ABEL [1]

Columbia University

AT THE present time through the agency of certain tests we are able to measure the eye-hand coordination of the young child at various stages of his mental development. The paper-folding tests of the Randall's Island Performance Series [2] and the Kuhlmann-Binet Scale [3] constitute a scale which, with the exception of the fifth year, is continuous from the second through the sixth year level. It was with a view toward filling in this gap that the present investigation was conducted.

Subjects: Since the object of the experiment was to find a paper-folding test which would measure the eye-hand coordination of the individual with a mental age of five years, it necessitated not only the use of children at this intelligence level, but also of those a year above and below.

The work was carried on in the Psychological Laboratory of the New York City Children's Hospital at Randall's Island. Here the children were selected according to their composite mental age which had previously been determined in the laboratory. On this basis then, excluding all epileptics and all those with motor disorders, eighteen cases with a mental age ranging from four years through four years, eleven months; nineteen cases from five years through five years, eleven months; and eleven cases from six years through six years, eleven months, were used. The chronological ages of the subjects ranged from six to eighteen years, with a range

[1] In collaboration with Dr. Louise E. Poull, New York City Children's Hospital, Randall's Island.

[2] The Randall's Island Performance Series, by Dr. Louise E. Poull, Columbia University Press, 1931.

[3] A Handbook of Mental Tests, by F. Kuhlmann, Warwick and York, Inc., 1922.

in IQ from 28 to 90. These last ranges were due to the fact that it was very difficult to obtain cases of the prerequisite mental age who at the same time were not suffering from motor disorders. Since, however, individual groups were homogeneous as far as the mental age was concerned, the scatter should not materially affect the results.

TABLE I

To show the ranges in chronological age, mental age, and IQ of the three groups

	No. of cases	Range of C.A.	Range of M.A.	Range of IQ
Group I	18	6 yrs. 3 mos.– 15 yrs. 3 mos.	4 yrs. 3 mos.– 4 yrs. 11 mos.	28-78
Group II	19	7 yrs. 0 mos.– 15 yrs. 8 mos.	5 yrs. 0 mos.– 5 yrs. 10 mos.	36-86
Group III	11	6 yrs. 11 mos.– 18 yrs. 1 mo.	6 yrs. 1 mo.– 6 yrs. 11 mos.	55-90

Procedure: The tentative paper-folding tests were divided into the following groups, each group of which consisted of four different tests: the first, those tests which consisted of two folds each and which were first demonstrated by the experimenter and then performed by the subject; the second group, those which consisted of three folds each, the examiner demonstrating the folds, one at a time, and the subject immediately performing each one; the third group, those tests which consisted of three folds each, the examiner first demonstrating the three consecutive folds and the subject then following; and lastly, the fourth group, in which each test consisted of four folds, the investigator demonstrating them one at a time and the subject immediately attempting each fold.

The following experimental procedure was then adopted: the subject was comfortably seated at a table facing the examiner and the paper-folding tests of the Randall's Island Performance Series and the Kuhlmann-Binet Scale given to set the upper and lower limits. The tentative paper-folding tests, sixteen in number, were next administered. In order to eliminate the effects of fatigue from the results, the test order was rotated with every individual.

In addition, upon any indication of restlessness or boredom the procedure was stopped and the child allowed to rest. The latter occurred but in one instance as the children were for the most part so interested in the "game" that they did not wish to stop playing it.

The scoring of the paper-folding tests of the Randall's Island Performance Series and of the Kuhlmann-Binet Scale was performed in the standardized manner described by their authors.[4, 5] In the tentative paper-folding tests the subject was given credit for having passed a test if he made no marked inaccuracy in the folds.

Results: The criterion, which was used to determine whether or not a given test measured the eye-hand coordination at the fifth year level, was the different degrees of difficulty the children of the three groups had in passing it. This in turn was measured by the percentage of children of each mental age group that passed it. The question then arose as to what percentage of the children with a mental age of five years should pass a test in order for it to be correctly placed at that mental level. It was finally decided that a test fulfilling the above criterion would be one which would be passed by 0 per cent of the group with a mental age of four years, by 75 per cent of the group with a mental age of five years, and by 100 per cent of the group with a mental age of six years.

When the results were finally tabulated, it was found that Test No. 1 very closely fulfilled the conditions demanded by the criterion set up, and that Test No. 2 approximated them, as shown in Table II. In the former test, the experimenter takes a piece of paper, six inches square, and instructs the subject to "watch carefully what I am going to do." He then folds the paper in half in such a way as to make a right triangle, and then folds it a second time as to make another right triangle, this time half the size of the first. The subject is then given a similar piece of paper and told, "Now you do that." If there are no marked inaccuracies in the folds, the subject is given credit for the test. In Test No. 2, the experimenter takes a paper, six inches square, and folding one

[4] Randall's Island Performance Series, by Dr. Louise E. Poull, p. 12.

[5] A Handbook of Mental Tests, by F. Kuhlmann, pp. 107-108.

corner towards the center, folds the opposite also towards the center, so that the two overlap. The paper is then placed on the table so that the overlapping is underneath. The subject is then told to do likewise. He is given credit for the test, provided the two folds are done correctly and provided that the opposite corners of the paper, when folded, overlap.

TABLE II

Showing the percentage of each mental-age group passing the paper-folding test of the Randall's Island Performance Series at the 4 year level, the Kuhlmann-Binet at the 6 year level, and the sixteen tentative tests for the 5 year level.

Group	Percentage Passing Test								
	R. I. Perf. Series	Kuhlmann Binet	Tentative Tests 1	2	3	4	5	6	7	8
I (M.A. of 4 yrs.)	100.0	0.0	16.6	38.8	11.1	5.5	5.5	0.0	0.0	0.0
II (M.A. of 5 yrs.)	100.0	10.5	73.6	78.9	26.3	57.8	36.8	68.4	31.5	0.0
III (M.A. of 6 yrs.)	100.0	100.0	100.0	100.0	72.8	100.0	81.8	81.8	63.6	45.4

Group	Percentage Passing Test									
	R. I. Perf. Series	Kuhlmann Binet	Tentative Tests 9	10	11	12	13	14	15	16
I (M.A. of 4 yrs.)	100.0	0.0	5.5	5.5	5.5	0.0	5.5	0.0	0.0	0.0
II (M.A. of 5 yrs.)	100.0	10.5	47.3	21.0	21.0	42.1	0.0	26.3	47.3	10.5
III (M.A. of 6 yrs.)	100.0	100.0	63.6	45.4	18.1	63.6	54.5	81.8	81.8	36.3

Conclusions: Due to the paucity of the number of cases and the fact that they were highly selected, all (with one exception) being subnormal in intelligence, the results obtained should not be regarded as strictly final. Test No. 1, however, may be looked upon as a tentative paper-folding test for the measurement of eye-hand coordination at the five year level, with Test No. 2 to be used as an alternate. The study should be repeated with normal subjects and the use of larger groups in order to insure more reliable results.

AUTOMOTIVE MANIPULATION IN RELATION TO VISION[1]

ALVHH R. LAUER and HAROLD L. KOTVIS

Iowa State College

PROBLEM

PREVIOUS studies of automobile driving (4) have suggested numerous factors of importance to safe manipulation of motor vehicles. One indication of this study was the importance of a large field of vision. Accident-prone drivers were found to have a constricted field. The normal eye is able to discriminate movement through an arc of about 150 degrees. By combining the two eyes through binocular vision approximately 180 degrees become visible in the combined, macular, paramacular, and peripheral fields.

While presentation of the stimulus is a prime requisite to visual discrimination in some degree, it does not follow that all objects falling within the visual field for a given time are perceived. Attention no doubt materially affects the perception of objects. Persons differ in the ability to perceive accurately at a glance. This does not seem to be a function of intelligence. A horse will react violently to a slight element in a total situation such as a piece of moving paper. The rider may not even notice such a slight element of the situation.

Since there are wide differences in the perceptual range of persons and since it seems related in some way to the successful manipulation of motor vehicles in traffic the present problem was set for experimental investigation. The problem may be set forth more explicitly by a series of questions. These may be stated as follows:

1. What is the relation between perceptual ability and visual acuity?

[1] Carried out under the auspices of the National Research Council Committee of "Psychology of the Highway" in cooperation with the American Optometric Association.

FIGURE I

Showing the driving apparatus as set up in the laboratory. The driver manipulates the small car (c) over the roadway (r) as the canvas landscape (l) rolls toward him. The speed of the canvas is controlled by the accelerator (a) which operates as a standard automobile. A shifting brush a.c. motor makes variation of the speed from 800 to 1750 r.p.m. very easy. The reduction in speed is taken care of by a regular transmission from an Essex car. Directions are given by means of the signs (s) which change with each revolution of the belt or canvas landscape. Other controls are the same as on a standard car. The score is recorded in terms of; (x) time taken for a trip of ten revolutions of the canvas, (y) errors made in negotiating turns, observing signs, etc., and (z) efficiency of the driver in keeping the small car on the roadway. This is registered in terms of contacts made. The better score is the greater number of contacts.

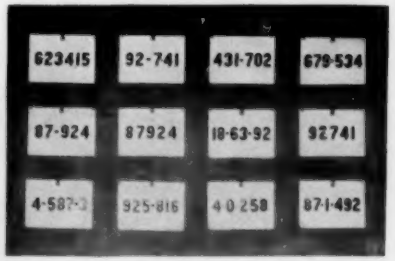

FIGURE II

Numbers used in the tachistoscopic study of perceptual span. The combinations were taken from automobile license plates. Number 623415 shown in the upper left was used to accustom the subjects to the procedure. The numbers are 1¼ inch Willson stock.

2. What is the relation between perceptual ability and size of the visual field?
3. Are there sex differences in perceptual ability?
4. Do younger persons observe more closely than older ones?
5. Is experience an aid to perceptual ability of a limited nature?
6. How is perceptual ability related to driving performance?
7. Are accident-prone individuals inferior in perceptual ability?

METHOD AND APPARATUS

In order to study the problem in its broader aspects a number of naive subjects were desired. The college student is likely to be aware of the conditions of an experiment and consequently is perhaps not representative of the general population. To secure such subjects required special arrangements since it is difficult to get any random sampling of the population to come into a psychological laboratory to serve as subjects.

Through the cooperation of the Iowa State Fair Association officials an experimental booth was established on the grounds at Des Moines during the State Fair. The walls were made of two-by-four studding covered on both sides by cornstalk board. An information desk was stationed in the front and small experimental booths were constructed and separated from each other by partitions and screens.

The booth was located in the Machinery Hall. This is a large steel structure with open sides housing the mechanical exhibits and various educational and commercial projects of the state. Certain incidental noises were more or less constant during the experiments and it is necessary to assume that these had equal effect upon all subjects. It may be of practical importance to measure such factors as perceptual span under these conditions since a driver actually performs on noisy streets and in heavy traffic where there are strong distractions and competing stimuli.

One experimenter at the information desk invited certain visitors to serve as subjects for the experiments. The object was to secure a cross-section of the population as nearly as possible. No children under twelve were taken as time would not permit and children can be studied at any place or time in the public schools. Various types of persons volunteered to take the tests, the results

of which they were given on a duplicate copy of the data sheet.

The booth was kept open five and one-half days during which time 339 subjects were studied. Four experimenters were busy all the time. The equipment and procedure used in the portable driving clinic were employed. A trained man was on duty all the time to make the visual examinations.[2]

The order of procedure was as follows: (1) Information secured by an interview relative to age, weight, height, occupation, and driving experience. The responses were checked for accuracy by cross-questioning. (2) Test for field of vision was made by the Nuchols chart. Stereoscopic vision was measured by a Keystone table model stereoscope. The Ishihara test was used for color-blindness. (3) Tachistoscopic tests of perceptual ability using a Whipple disc tachistoscope. (4) Complete visual examination including a check of visual acuity, the phorias, and astigmatism. The Clason Acuity Meter was used. A test of glare was also made by the Glarometer.[3] (5) Driving performance was measured on the indoor driving apparatus shown in Figure I. (6) Other measurements not considered in this study were made.

When through the experiments each subject received a duplicate sheet showing his record. Suggestions as to possible means of improving his driving were given. This was found to be an excellent motivating scheme for adult subjects. If they feel they are getting something in return for their efforts, better rapport is established. After leaving the clinic many would send their friends in to be examined. The psychologist at the information desk made all interpretations of the records after they were completed. In this way the subject was led to a better appreciation of the nature and use of the tests.

It will not be necessary to describe all the procedures in detail. The data were recorded upon a special mimeographed sheet devised for the purpose. A place was allowed for every measure-

[2] The optical measurements were made through cooperation of the Iowa State Association of Optometrists by courtesy of Dr. V. V. Kirby, President of the Association.

[3] Made by the Glarometer Co., Chicago, Ill.

ment and in the order which they were given. At the end of the sheet a number of explanations were given about the various tests and their application to automobile driving.

The driving apparatus layout has already been described under Figure I. The subject is placed at the wheel and told to manipulate a miniature car (c) over the moving roadway (r) directly ahead of him. A complete set of automatic controls keeps the trip to ten revolutions of the canvas highway. Time, efficiency in keeping on the highway, and errors are automatically recorded by counters. These items were used as variables in the statistical analysis.

The Nuchols test is merely a paper protractor cut in semi-circular form about 14 inches in diameter and marked off in degrees. A small hole at the center allows the device to fit up closely around the nose on a level with the eye. The experimenter sits in front of the subject, fixates his eyes, and moves a target at one side or the other to check the perception of movement. Beginning out of range he moves in until movement is reported. A record of the angle of declination from the median plane is read off directly from the protractor. The measurement is subject to an experimental error of about 5°.

A standard Keystone stereoscope with illumination for slides was used. Slides number PG5 and PG6 were employed to measure acuity. Stereoscopic vision was checked by slides 033 and 034 F. Acuity only will be treated in this paper.

A form was provided for the Ishihara tests and the actual numbers missed were tabulated. The test of perceptual ability is described in the next paragraphs.

MEASURING PERCEPTUAL SPAN

The apparatus consisted of a Whipple tachistoscope and a set of twelve exposure cards, shown in Figure II. The tachistoscope was modified somewhat from that as devised by Whipple. The opening in the disc was set to measure 103 mm. by 223 mm. and only one bob or weight was used on the pendulum, this being set at a reading of 30.95. In addition the arrangement of pulleys was eliminated and the cards placed in the holder directly in line with the front aperture. The shutter was opened to a distance of 15.18 cm. at the outer edge. The time of exposure with these ad-

justments was calibrated by an Ewald chronoscope and found to have a mean span of 105 sigma with an error of ±5 sigma. The light conditions were measured by use of a Weston Illuminometer. The eye was adapted to from three to four foot candles of light, the amount of illumination in the room in which the test was given. The tachistoscope was furnished with a 25-watt clear glass showcase Mazda bulb. An average reflection factor of 13.4 foot candles at the middle of the opening in the face of the tachistoscope and an average direct light on the exposure cards at their center of 27.1 foot candles was maintained.

The cards consist of black digits 4.5 cm. in height and 3 cm. wide placed in groups of five or six on white cards measuring 14.6 by 22.2 cm. The numbers were spaced at about 4.5 cm. from center to center. In some cases dashes were included in the grouping. See photograph below. The first card in the upper left hand corner was used to accustom the subject to the method used. The others were given as described in the text.

The directions given the subject were as follows, "In this test you will be shown groups of numbers. Try to read as many as you can. This first card will not count but will be given for practice." (First card shown.) "What did you see? Now we will begin." The remaining cards were exposed and the results recorded by the experimenter after each. In most cases six cards were used. The first and last six of the eleven being used separately. All eleven were given to a part of the subjects.

Visual examinations were carried out under a simplified form of the standard clinical (1) procedure as used by optometrists. One form of this abbreviated procedure has been described in a previous publication (4, page 90). A more recent form is now being completed and will be described later.

The susceptibility to glare was measured by the Glarometer and recorded in terms of scale readings on this instrument. A rheostat placed in series with the 250-watt Mazda bulb which is mounted behind a ground glass and a condensing lens which distributes the light. As the current increases, the reading matter placed upon opaque paper attached to the ground glass becomes invisible through the icreased intensity of the background illumination. Subjects vary between the limits of scale readings as low as 25 up

to readings of 100. This seems to be a function of the adaptation of the eye but it is relatively consistent. Readings on bright days are usually higher than those made on dark days. One weakness of the test is that the quality of the light or characteristic wave length differs with the amount of current.

The glare difference measurement was made by comparing readings with the naked eye and those made when protected by a Calobar C tinted lens. The measurements were made first without filter, second with filter, and third without filter. The first and third measurements were averaged and compared with the second measurement. In this way fatigue or adaptation effects were thought to be mostly eliminated. Many subjects can stand more light without being blinded when a tinted lens of the type used is held before the eye but this is not universally true.

The speed of movement measurements were made by means of a telegraph key with conditions approximately the same as used by Ream (2) except that a ten second tapping period was used instead of five seconds. Five trials were made with the preferential hand and the average of these used as the score. Strength of grip scores were made by averaging two right and two left hand trials on a standard dynamometer.

At the completion of the tests only 188 subjects of the total of 339 could be used for this particular study. The data were coded and placed upon Hollerith cards. The correlations given below in Table I were calculated by the method of Wallace and Snedecor (3) using the so-called Ayres formula:

$$r = \frac{\Sigma AX - (\Sigma A)Mx}{\sqrt{\Sigma A^2 - (\Sigma A)M_A} \cdot \sqrt{\Sigma X^2 - (\Sigma X)Mx}}$$

(*Legend* for Table I)

Ag=age.
Ex=experience in terms of years driving.
Wt=weight of driver.
Ht=height of driver.
Mpr=average driving speed.
Cd=number of cars driven by the driver.
Vf=extent of the visual field in degrees.
As=acuity measured by the stereoscope.
Dc=driving score in contacts made.
De=driving score in errors made.
Dt=driving score in time to complete the trip.
Ao=acuity as measured by optometric methods.
Gi=glare index. Scale readings on Glarometer.
Gd=glare difference. With and without filter.
Pn=per cent of numbers correctly perceived. Perceptual span.
Ai=astigmatic index.
St=strength measured by the hand dynamometer.
Sm=speed of movement. The tapping test.

TABLE I

Correlation of Factors Relating to Automobile Driving and Perceptual Ability

	Ex	Wt	Ht	Mpr	Cd	Vf	As	Dc	De	Dt	Ao	Gi	Gd	Pn	Ai	St	Sm
Ag	+.6652	+.3098	−.1454	−.1752	+.0100	−.0435	−.2095	−.0644	+.2240	+.2456	−.2481	−.2765	+.0361	−.1655	+.0703	−.0800	+.1219
Ex		+.3220	−.0270	+.0667	+.2475	+.0294	−.0345	−.1172	−.0158	+.1738	−.1090	−.1271	−.0631	−.0693	+.0856	+.1179	+.2594
Wt			+.4673	+.1725	+.1433	+.0237	−.0686	−.0398	−.0293	+.1455	−.0744	−.0165	+.1477	−.0314	+.1418	+.5001	+.1647
Ht				+.2155	+.1385	−.0238	−.0079	+.0644	−.2349	−.0193	+.0232	+.1781	−.0430	+.1086	+.1600	+.6340	+.1384
Mpr					+.1987	+.1412	+.0858	−.0867	−.2216	−.0230	+.3093	+.0694	−.0093	+.2076	−.0186	+.3434	+.0916
Cd						+.1504	−.0458	−.0092	−.1508	−.0198	+.1721	+.1544	+.0688	+.1317	−.0129	+.1414	+.2408
Vf							+.0414	−.0851	+.0291	−.0267	+.0690	+.0647	−.0374	+.0165	+.0514	+.0155	+.0742
As								+.0512	−.1781	−.1271	+.2570	+.2364	−.0232	+.2054	−.0077	+.0658	−.0471
Dc									−.0897	−.1746	+.1883	+.2632	+.0759	+.0050	+.0486	+.0663	+.0656
De										+.1334	−.0690	−.2064	−.0281	−.2059	+.0301	−.2333	−.1196
Dt											−.0561	−.2437	−.1078	−.0519	+.1474	+.0263	+.0662
Ao												+.2664	−.0411	−.0366	−.1302	+.1177	−.0488
Gi													+.0455	+.1225	−.0797	+.1424	+.0015
Gd														−.0043	+.0268	+.0371	−.1140
Pn															+.0154	+.0566	+.0300
Ai																+.0969	+.0513
St																	+.2056

N=188. The P.E. range has been computed and shows any correlation over .14 to be significant. Certain coefficients are not closely related to our study but it was easier to calculate and include, than to omit them.

(*Legend* on page 427)

SUMMARY AND CONCLUSIONS

Table I shows the following results to be significant.

A. Those directly relating to the problem attacked.

1. Acuity measured by the stereoscope correlates +.2054 with perceptual ability. It correlates −.0366 with acuity measured by optometric techniques. The two measurements of acuity correlate +.2570. This indicates a difference in the two types of measurements and the possible operation of factors related to stereoscopic vision and perceptual ability since stereoscopic vision was introduced into the measurement of acuity. The letters were large enough to be visible to one with five per cent vision at the distance exposed. The angle subtended by the total number in the horizontal plane would be about 8° 19′. Acuity would drop off somewhat at this distance from the macula.

2. There is near zero correlation between visual field and perceptual ability.

3. Experience as measured in years driving does not seem to be correlated with perceptual ability. When the number of cars driven is used as a criterion a barely significant positive correlation exists. However the factor of age is not considered in either case. It will be discussed below.

4. Driving errors, while actually performing, showed a negative correlation with observational ability. This suggests an advantage for the person who can observe accurately when in a tight place. There seems to be no relation between manipulative ability and the power of observation. The negative correlation of −.1655 between the percentage of numbers read and motility might suggest a slight advantage for the person who tends to move slowly.

5. Observational ability correlates somewhat with average driving speed, r=+.2076, which seems to indicate a possible tendency for fast drivers to be more observant. Also the astigmatic index correlates with observational ability +.0154 and indicates no tendency to compensate for visual deficiencies.

B. Factors indirectly relating to the problem attacked.

6. Persons with better vision are affected less by the glare of a strong light. This is shown by the correlations of +.2364

(Legend on page 427)

between acuity as measured by the stereoscopic slides and glare. The correlation of glare with acuity as determined by optometric methods gave an r of +.2664, which further substantiates this observation.

7. Older drivers tend to drive slower, make more minor errors, are affected more by glare, and do not observe as well as younger drivers.

8. The more experienced drivers are inclined to overlook unessentials, to take more time on the driving test, to be more affected by glare, and to have superior motility.

9. Fast drivers, on the average, have driven more types of automobiles, have a larger visual field, observe more closely, have superior vision, are stronger and tend to have higher motility rate.

10. Visual field correlates almost zero with everything but driving speed and the amount of experience with different types of cars. These coefficients are barely significant.

11. The driving performance correlates highest with age, experience and susceptibility to glare. The person with astigmatism does not compensate to any extent, is slightly older, and obviously has impaired acuity.

12. Persons with low visual acuity are older, drive slower, have driven fewer types of cars, observe fewer details when driving, and are not as strong physically.

13. No significant difference was found between accident and non-accident drivers on perceptual ability. No sex differences in perception were found in this study.

SUMMARY

There is some evidence that accident drivers on the average have higher motility rate, are stronger, live in large towns and tend to compensate. They do not show superiority on observational ability. Compensation is highly significant since it will be necessary to partial out certain factors in the final analysis of accident causation. The results presented in this paper are offered only as a representation of 188 cases studied during one week under the conditions described above. The correlation table is presented as typical of these results.

About 50 per cent of the relationships shown here are not linear

functions which may explain the low correlation coefficients.

A more complete study including a larger number of cases is being worked out at this time. The present research was aimed primarily at the problem of perception. The general conclusion is that perceptual span is a complex factor and is only slightly related to such specific functions as driving an automobile. It is more closely related to age and habituated response patterns than to basic capacities such as acuity of vision and size of the visual field. A number of indirect relationships brought about by compensations tend to mask many significant factors. These effects will need to be considered in any interpretative review.

REFERENCES

1. Ellis, C. R., Hare, R. A., and Lauer, A. R., The Prevalence of Visual Defects and their Relation to Automobile Driving. 1932. Jour. Amer. Optometric Assoc., *7*, No. 9, 1-7.

2. Ream, M. J., The Tapping Test; A Measure of Motility. 1922. Psychol. Monog., *31*, 293-319.

3. Wallace, H. A., and Snedecor, G. W., Correlation and Machine Calculation. 1925. Iowa State College Bulletin, *30*, No. 4, 5-11.

4. Weiss, A. P., and Lauer, A. R., The Psychological Principles in Automotive Driving. The Ohio State University Press. 1931. Contributions to Psychology No. 11.

AN EVALUATION OF CERTAIN INTELLIGENCE TESTS

M. V. SEAGOE

Director of Research, Garvey School District

THE relative merit of different group measures of intelligence is a problem constantly before educators and research workers. The need for an evaluation of such tests has been partially met by the coefficients of reliability and validity established for the newer measures, but more elaborate comparison between tests is needed.

The subjects included in the present study are drawn from the elementary schools of the Garvey District in Los Angeles County. The populations for the nine comparisons varied from 61 to 256, averaging 130. The pupils are representative; the community is over 97 per cent white American, consisting of skilled and semi-skilled workers. The groups used in this study are selected only in permanence of residence; in order to take two tests under the procedure used the child must have resided in the district an average of two years. This factor does not invalidate the comparisons, for in each case the population is constant.

The data are taken from intelligence test results over a period of five years. The original survey covered all grades; thereafter tests were given each semester in 1B, 3B, 5B and 7B. The tests used were the Detroit First Grade Intelligence Test in 1B, the Detroit Primary Intelligence Test in 3B, the National Intelligence Test Scales A and B in 5B, and the Terman Group Test of Mental Ability in 7B. In the original survey Haggerty Delta I and the Pintner-Cunningham Primary Mental Test were used in primary grades to supplement the above tests.

The average lapse of time between first and second tests was two years. It may possibly be argued that this situation will introduce an age variable which will vitiate the results. Such an effect would probably be negligible for three reasons. (1) For a fairly large normal group such as that used it should be safe to assume

normal growth. There should be no factor tending to bias the whole group, though individual differences in growth might tend to lower all coefficients of correlation. (2) Each test was used near the optimum age for which it was designed, and not toward one end of the scale where standards are less reliable; such spacing should make the comparisons more rather than less valid. (3) Such spacing represents the normal practice in the use of intelligence tests.

The data tabulated included chronological ages at the time of testing and intelligence quotients derived. Statistical work included computation of the mean, standard deviation and probable error of the mean for each test for each group, the difference between the means and its probable error, the coefficient of correlation for each pair and its probable error, and the coefficient corrected for attenuation when the reliability of the test could be ascertained.

The coefficients might have been higher had it been possible to use raw scores instead of derived figures, but only records of intelligence quotients were available. This condition affects all tests equally, however, and does not invalidate comparisons between coefficients obtained on such a basis.

The results are summarized in Table I. The column headings are self-explanatory, with the possible exception of the last two which indicate the coefficient of reliability and the coefficient of correlation corrected for attenuation.

The significant points are two. (1) In the first four pairs of tests the difference between mean intelligence quotients derived for the same children on the two tests is statistically significant. (2) In the first four comparisons, involving strictly primary tests, the coefficients of correlation range somewhat lower than those for the last five pairs, which do not involve strictly non-verbal tests.

The tests studied appear to fall roughly into two divisions: (1) Performance tests for first and second grades, such as the Detroit First Grade Intelligence Test and the Pintner-Cunningham Primary Mental Test. These measures yield intelligence quotients from five to eleven points higher than those with which they were compared, and fail to correlate as highly with other tests as do tests involving a larger verbal element. (2) Middle and upper

TABLE I

Comparison of Intelligence Tests

Test	N	Chron. Age Mean & Sigma		Intelligence Quotient Mean	Sigma	Prob. E.M.	Diff. Means	r. P.E.	Reliab.	Atten.
1. Detroit First Grade	120	6–4	5.7	114.1	13.3	0.82	* 10.3 ± 1.2	.642 ± .036	....	....
Detroit Primary		8–8	8.1	103.8	13.5	0.83			.93	
2. Detroit First Grade	71	6–3	4.7	109.6	12.8	1.02	* 6.3 ± 1.3	.655 ± .046	....	....
Haggerty Delta		8–8	5.8	103.3	10.0	0.80			....	
3. Pintner-Cunningham	186	7–7	7.4	110.3	12.6	0.62	* 5.1 ± 0.8	.587 ± .032	....	....
Detroit Primary		8–6	7.9	105.2	11.8	0.58			.93	
4. Pintner-Cunningham	96	7–6	7.3	111.1	12.5	0.86	* 11.2 ± 1.3	.519 ± .051	....	....
National B		10–10	6.5	99.9	14.0	0.97			.949	
5. Detroit Primary	57	9–0	12.2	107.2	12.5	1.11	4.6 ± 1.7	.700 ± .046	.93	.756
National A		10–3	6.7	102.6	14.0	1.25			.922	
6. Detroit Primary	187	8–9	9.5	107.2	14.5	0.71	3.6 ± 1.1	.727 ± .023	.93	.774
National B		10–8	8.2	103.6	16.0	0.78			.949	
7. Detroit Primary	61	9–4	11.5	107.1	16.1	1.38	0.6 ± 1.8	.765 ± .036	.93	.841
Terman		12–4	9.8	107.7	14.0	1.20			.89	
8. National A	256	10–4	11.8	109.6	17.0	0.72	0.3 ± 0.8	.797 ± .015	.922	.881
Terman		12–5	8.7	109.9	12.0	0.50			.89	
9. National B	137	10–7	9.9	106.9	16.4	0.83	0.8 ± 1.1	.869 ± .014	.949	.945
Terman		12–6	8.4	107.7	11.8	0.68			.89	

* The difference between the means is statistically reliable.

grade tests, such as the Terman Group Test of Mental Ability, the National Intelligence Scales, and the Detroit Primary Intelligence Test. These tests appear to yield comparable intelligence quotients and to correlate more highly with each other. The Terman and National tests agree most highly; the Detroit Primary Intelligence Test agrees with them less closely, perhaps because it represents a transition from the performance to the verbal type of intelligence test.

The difference resolves itself largely into one between primary tests which approach the problem through oral directions, pictures, drawings, and such less verbal problems, and upper grade tests which lean more largely on problems involving some school achievement. The primary and upper grade tests, through their difference of approach, seem to define intelligence somewhat differently. In this connection, Kelley's thesis that ". . . some 90 per cent of a general intelligence test and an all-round achievement test measure the same thing," [1] and ". . . most of the distinctions drawn between intelligence and achievement are spurious" [2] is of interest. The measures used by Kelley were for the most part upper grade tests; it would be illuminating to see whether the relationship holds for primary intelligence tests and primary achievement batteries. With such close agreement between upper grade achievement and intelligence tests, it may be argued that intelligence as measured by an upper grade test is more closely related to achievement than to "intelligence" as it is defined in primary performance tests.

In summary, a comparison of intelligence quotients on different group tests given at two-year intervals shows: (1) The Detroit First Grade Intelligence Test and the Pintner-Cunningham Primary Mental Test apparently yield intelligence quotients higher than other tests. (2) Early primary intelligence tests of a performance type correlate with middle and upper grade tests .50 to .65; middle and upper grade tests correlate with each other .70 to .87, or .75 to .94 corrected for attenuation.

---

[1] Kelley, Truman L. *Interpretation of Educational Measurements*, p. 193. New York, World Book Company, 1927.

[2] *Ibid.*, p. 209.

These facts appear to indicate some difference in the attribute measured by primary and upper grade intelligence tests. Primary tests will agree less with subsequent upper grade tests than upper grade tests will agree with themselves, and will probably correlate less highly with school achievement than will upper grade tests. Such facts should be considered in interpreting test results.

A COMPARISON OF THE HANDWRITING OF PUPILS IN CERTAIN ELEMENTARY SCHOOL GRADES "NOW AND YESTERDAY" [1]

HARRY TENWOLDE

Union Grammar School, Imperial, California

IT IS the purpose of this short study to consider some data pertaining to the problem of excellence of penmanship quality in the public schools now and in certain former decades. The data as to handwriting quality in the past decades (specifically, in 1879 and in 1912) were obtained in a report prepared by an earlier research worker.[2] The data with respect to modern public school penmanship quality were secured by the present writer during the second semester of 1930-1931 in the El Centro, California, public schools.

Data are considered in this discussion for four elementary school grades, grades five through eight. Since Bole used the Thorndike Handwriting Scale as his criterion for securing measures of the quality of the penmanship of public school pupils in 1879 and in 1912, the writer necessarily used it also. The procedure of the study involved only the scoring of the samples of 1931 pupil penmanship, forty pupils in each grade group, and the statistical computations for the three groups in 1879, 1912, and 1931. The findings are considered hereafter for each grade separately.

Grade five. Twenty-four pupils, nine boys and fifteen girls are included in the 1879 pupil group. The penmanship quality scores for this group ranged between 9.0 and 13.0, with a mean of 11.04 and a standard deviation of 1.14. The scores for the thirty-seven

[1] This study is based on part of a Master's thesis prepared under the direction of Dr. A. S. Raubenheimer of the University of Southern California.

[2] Bole, S. J. Penmanship in Grades V, VI, VII, and VIII of the Urbana Public Schools. Unpublished Master's Thesis, 1912, 58 pages. Urbana: The Library of the University of Illinois.

pupils in the 1912 group ranged between 11.0 and 14.0. The mean for this group was 12.65; the standard deviation, 0.81. The 1912 group of pupils included seven boys and thirty girls. The 1931 pupil group included forty subjects, twenty boys and twenty girls. The penmanship quality score range included scores from 8.0 to 12.0. The mean was 9.60 and the standard deviation was 0.89. The pupils in the three groups in penmanship quality excellence indicate that the handwriting of 1912 ranks first; that of 1879 second; and that of 1931 a weak third. The significance of the differences is discussed in Table I.

The data in Table I indicate that the handwriting of pupils in grade five of the elementary school in 1931 is significantly weaker than that of pupils in similar school situations in 1879 and in 1912. It is not possible to make positive statements because the groups considered are quite small. In addition, in each instance only one school system is considered and the pupils of this given school system may not be a typical sampling of the total school population in the given school year.

TABLE I

Measures of the Reliability of Differences Between Mean Scores, Grade Five

Groups compared	Ratio of Diff. to S.D. (Diff.)	Percent of Diff./S.D. of Diff. necessary always to insure a difference	Chances in 100	Observed difference favors
1897–1912	5.99	100.0	100.0	1912
1897–1931	5.27	100.0	100.0	1879
1912–1931	15.72	100.0	100.0	1912

Grade six. Only ten students are included in the 1879 grade six group, three boys and seven girls. The score range included quality scores from 11.0 to 15.0 inclusive. The mean score was 12.5, and the standard deviation, 1.20. In the 1912 grade six group eight boys and sixteen girls are found, twenty-four pupils in all. The mean score for this group was 13.13. The standard deviation was 0.60. The range included scores between 12.0 and 14.0.

The 1931 grade six group included forty pupils, twenty boys and twenty girls. The scores of this group of forty pupils ranged between 7.0 and 11.0 inclusive, with a mean of 9.10, and a standard deviation of 1.0.

As in grade five, the penmanship of grade six pupils in 1931 is much inferior to that of both 1879 and 1912. The statistical significance of the difference is considered in Table II. There is no statistically significant difference in penmanship quality in grade six between 1879 and 1912, but the differences in penmanship quality in this grade in 1879 and 1931 and between 1912 and 1931 are statistically significant. The pupil groups in 1879 and 1912 are so small that definite conclusions cannot be drawn.[3]

TABLE II

Measures of the Reliability of Differences Between Mean Scores, Grade Six

Groups compared	Ratio of Diff. to S.D. (Diff.)	Percent of Diff./S.D. of Diff. necessary always to insure a difference	Chances in 100	Observed difference favors
1879–1912	1.58	52.7	94.0	1912
1879–1931	8.27	100.0	100.0	1879
1912–1931	20.15	100.0	100.0	1912

Grade seven. The 1879 pupils in grade seven total sixteen. The group included four boys and twelve girls. The penmanship quality scores for this group ranged between 12.0 and 15.0. The mean score was 13.06; the standard deviation, 1.09. Six boys and twenty-four girls were included in the 1912 grade seven group, a total of thirty pupils. The penmanship quality score range included scores between 12.0 and 15.0. The mean score was 13.37 and the standard deviation was 0.83. Forty pupils are included in the 1931 group, twenty boys and twenty girls. The scores for

[3] See "limitations to reliability formulas" in Garrett, H. E. Statistics in Psychology and Education, New York: Longmans, Green and Company, 1926. pages 128-133.

this group ranged between 7.0 and 12.0, with a mean of 9.75 and a standard deviation of 1.32. The data are treated further in Table III.

TABLE III

Measures of the Reliability of Differences Between Mean Scores, Grade Seven

Groups compared	Ratio of Diff. to S.D. (Diff.)	Percent of Diff./S.D. of Diff. necessary always to insure a difference	Chances in 100	Observed difference favors
1879–1912	1.00	33.3	84.0	1912
1879–1931	9.68	100.0	100.0	1879
1912–1931	13.71	100.0	100.0	1912

As in grades five and six, the penmanship quality of the 1931 grade seven pupil group is markedly below that of pupils in a similar school situation in 1879 and in 1912. As in grade six, the difference in penmanship quality between 1879 and 1912 are not statistically significant, but the differences between 1879 and 1931 and 1912 and 1931 are statistically significant. The pupil group in 1879 is small. As in grade five and in grade six, it is difficult to draw final and definite conclusions on the basis of findings in a single group.

Grade eight. Seven boys and nine girls, sixteen pupils in all are included in the 1879 grade eight group. The score range for this group included scores of 12.0 through 16.0 inclusive. The mean score was 14.06, and the standard deviation was 1.35. The 1912 grade eight group included thirty-two pupils, six boys and twenty-six girls. The mean score for this group was 13.53. The standard deviation was 0.75. The scores for this group ranged between 12.0 and 15.0. The forty pupils in the 1931 group in grade eight included twenty boys and twenty girls. The penmanship quality score range included scores between 7.0 and 14.0. The mean score was 10.6 and the standard deviation was 1.46. Table IV presents a further evaluation of these data.

As in the instances of grades five, six, and seven the penman-

ship quality of 1931 in this grade is markedly below that of 1879 and 1912. Again, as in grades six and seven, the small numbers of cases in the 1879 group makes the interpretation of some of the findings difficult.

TABLE IV

Measures of the Reliability of Differences Between Mean Scores, Grade Eight

Groups compared	Ratio of Diff. to S.D. (Diff.)	Percent of Diff./S.D. of Diff. necessary always to insure a difference	Chances in 100	Observed difference favors
1879–1912	4.61	100.0	100.0	1879
1879–1931	8.46	100.0	100.0	1879
1912–1931	11.02	100.0	100.0	1912

Summary and conclusions. The problem considered in this study is whether or not the penmanship of pupils in grades five through eight inclusive is inferior to, equal to, or superior to that of pupils in similar school situations in 1879 and in 1912. The findings are difficult of interpretation in some instances because of the small number of cases involved in given grades, particularly in the 1879 groups. Since the number of cases is small in each of the grade groups considered and since the sex sampling in 1879 and in 1912 is uniform in no one group it is obvious that the assumption underlying the interpretation of reliability formulas involving the standard error, i.e., that the measures are distributed in accordance with the normal probability curve does not hold. The findings may be affected by other limitations, of course, but this is the most obvious and most serious one.[4] On the other hand, the fact remains that in four grades the penmanship of 1931 pupils is always inferior to that of pupils in the same grades in 1879 and in 1912. The differences are always statistically significant. This unanimity in the findings for the four grades does much to offset the weaknesses which affect the interpretation for one given grade. If we may assume that the pupils in the single 1931 school situation considered are a typical sample of such

[4]Garrett, H. E., loc. cit., pages 142-145.

pupils, then we may safely assert that pupils in public school grades five through eight in 1931 write with a penmanship quality which is significantly poorer than that of pupils in public schools in 1879 and in 1912.

A STUDY OF VOCATIONAL ABILITIES OF STUDENTS WHO HAVE ATTENDED THE ILLINOIS SCHOOL FOR THE DEAF[1]

VERNE WESLEY LYON

Illinois Institute for Juvenile Research, Chicago

THE purposes of this study may be briefly stated as follows: first, to determine so far as possible the vocational activities of former students at the Illinois School for the Deaf during the past ten years; second, to gain some insight into the problems that confront the deaf in their various attempts at vocational adjustment; third, to determine to what extent the vocational training at the school has been helpful in the economic and social adjustment of the deaf child; and fourth, as a result of this study to suggest possible revisions in the present system of vocational training.

As a preliminary step in the study of the vocational adjustment of the deaf, a letter was sent to various state schools for the deaf in representative parts of the United States and Canada.

The purpose of this letter was to determine first whether the same procedure is adopted generally in the vocational training programs for the deaf, and second, to determine any new or improved methods which might be employed in facilitating the vocational adjustment of the deaf.

Nine schools cooperated in the work of this study. The schools replying were: California School for the Deaf, Columbia Institution for the Deaf (Washington, D. C.), Indiana State School for the Deaf, Clark School for the Deaf, Minnesota State School for the Deaf, Missouri School for the Deaf, Institute for the Improved Instruction of Deaf Mutes (N. Y.), Pennsylvania Institute for the Deaf, and Ontario School for the Deaf.

The replies from these schools present a representative sample

[1] Studies from the Institute for Juvenile Research, Chicago. Series C, No. 234.

of the accepted methods and procedures adopted in the vocational training of the deaf.

Questions were formulated with the purpose of obtaining information about the most important factors in the vocational training program. It was found that:

1. Eight schools begin definite vocational training for boys between the ages of thirteen and fourteen; one school begins training at the age of twelve.

2. In general, the four most common courses taught are (1) printing, (2) cabinet making and carpentry, (3) shoemaking and repairing, and (4) baking. Some schools, however, are considerably more versatile than others.

3. On the question, "What method do you use for determining the course that is best adapted to each boy?" there appeared to be the most dissimilarity of opinion and approach. Since this factor is of vital importance in the interpretation and understanding of the vocational adjustment problem, special reference will be made to it in the latter part of this study.

4. The average number of hours in the school day was found to be seven, and the average time devoted to vocational training was between two and three hours a day.

In general it was found that on the four points just discussed, with the exception of Number 3, there is in these schools a great similarity of procedure, a widespread adoption of common factors that can be closely compared with, and typified by, the vocational training program at the Illinois School for the Deaf. This fact is considered to be significant because the general findings and recommendations to be made as a result of this study at the Illinois School for the Deaf may be said to be applicable in some degree to other similar institutions.

The most outstanding contribution designed to facilitate the vocational adjustment of the deaf was received from Pennsylvania Institution for the Deaf at Mt. Airy, Philadelphia. Under the supervision of Harry B. Brown, principal of the vocational school, a comprehensive survey was made of the vocational activities of its former pupils during the past twelve years. The medium through which this survey was made was a questionnaire which produced such satisfactory results that with a few minor revisions

it was used also in the study made at the Illinois School for the Deaf.

Since it would seem that those deaf students who have been trained under the present system and who have had an opportunity to apply this training in a competitive industrial field would be in the best position to offer suggestions and criticisms of the vocational training program as it exists today, the questionnaire mentioned above was sent to all former students of the Illinois School for the Deaf who had either graduated or left school during the past ten years and who had received at least two years vocational training while in school.

The total number of questionnaires sent to former pupils was three hundred forty. Replies were received from one hundred twenty-five persons, or thirty-seven per cent. Of these, forty-six were from graduates while seventy-nine were from students who had left school before graduation.

The average time these pupils had spent in school was nine and one-half years, and the average time elapsed since leaving school was five years.

The ages of the pupils vary from seventeen to thirty-five years. The average age is between twenty-four and twenty-five years. Of the one hundred twenty-five replying, sixteen per cent are married and forty-seven per cent report that they are self-supporting.

One hundred seventy-one questionnaires were sent to boys and sixty-nine replies were received. One out of every four boys studied more than one trade or vocation during his stay in school. Sixty-two per cent did not follow the trade they had acquired in school. Of the remaining thirty-eight per cent who did follow their trade, the period of employment varied from one month to six years, indicating some degree of instability. At the present time sixty-six per cent report that they are working, yet only five boys or eight per cent are following the trade acquired at school. Many reasons are offered for the changing of trades, the most common one being the inability to find work at the acquired trade. The interesting fact is, however, that sixty-six per cent are working, indicating that the deaf are capable of assuming industrial responsibility. Why is it, then, that only eight per

cent follow the trade acquired at school? There must be some answer to this question.

Seventy-seven per cent suggest more time for trade work in school while three per cent are opposed to the extension of time.

One hundred sixty-nine questionnaires were sent to girls. Fifty-six replied. Each girl studied two or more trades during her stay in school. Thirty-six per cent report that they are working, yet only sixteen per cent are following the trade acquired at school. Eighteen per cent are married and fifty per cent report that they are housekeeping. Sixty-nine per cent suggest additional time for trade learning while thirty-one per cent oppose it.

The interpretation of the results of this study are deductions based entirely on the information supplied by the deaf. For clarity in presentation, the boys and girls will be discussed separately.

INTERPRETATION OF RESULTS

VOCATIONAL TRAINING FOR DEAF BOYS

It has already been noted that on the basis of the information received only eight per cent of the boys who have left school during the past ten years are working at the trade they acquired in school. This rather startling fact presents two problems for consideration. First, does the relatively small number of boys who are ultimately able to profit by vocational training in school warrant the large expenditure of time, effort and money that is now being spent on it? Or is it possible to direct these forces in industrial channels which would yield more benefit to both the individuals and the school?

It seems plausible to assume that the purpose of vocational training in general should be to develop proficiency in some specific trade that will enable an individual to compete on more equal terms with other individuals.

In the case of a deaf child the need for training in some specific trade is of even greater importance, since the degree of proficiency would appear to be the significant factor in the competition between deaf and hearing individuals. School offers practically the only opportunity for the development of abilities. It requires a great deal of patience, understanding and time to instruct prop-

erly a deaf person. This is a difficult task, even for those who are specially trained and qualified to do it. Consequently, the possibility of a deaf person receiving adequate instruction and training after leaving school is somewhat remote because of his physical handicap and because there are so few facilities for the training of the deaf in rural communities or small towns to which most of these children return.

Therefore it is obvious that the deaf are dependent upon school training for their success in trade work and the fact that so few are now profiting as a result of this training does not detract from its importance, but rather emphasizes the need for the adoption of improved methods in instruction.

Sixty-six per cent of the boys are working, indicating that there are positions in which the deaf are capable of adjusting. The reasons then for the failure to adjust in those vocations acquired in school are presumably not due to the inability of the deaf as a group to assume industrial responsibility.

The deaf have presented many practical suggestions for the adoption of new trades in keeping with industrial progress. An important factor involved in the selection of new trades is the territorial location of the school and the predominating communities from which deaf children are assembled.

Thirty-one trade suggestions were made, the most frequent being tailoring, barbering, auto mechanics and farming. The adoption of these four trades for vocational training would not require any prohibitive increase in equipment and might well replace some of the less profitable trades now being taught. Tailoring, barbering, and auto mechanics are staple trades at which the deaf should have no difficulty in adjusting satisfactorily, and since the sources of many children at the Illinois School for the Deaf are rural communities, instruction in farming would be very helpful for those so inclined.

Assuming that a deaf child has been properly trained in some vocation and is ready to leave school, the next important problem is finding a job. The burden of this responsibility should not fall directly on the deaf individual because of his inability to converse and to expound his abilities to a prospective employer. There is likely to be some skepticism on the part of the employ-

ing agent as to the type of training and the degree of proficiency that a deaf child may possess. He must be convinced that the deaf are capable of competing industrially with hearing individuals and for this purpose it seems necessary to have an intermediary.

The establishment of field workers to contact the community from which a deaf child comes and to which he is likely to return, in order to determine the trade possibilities of that community and to assist in securing employment, would be a remedial approach to this problem.

It has long been thought that the trade possibilities for the deaf were extremely limited and for many years the deaf child has been very restricted in his selection of a vocation due to this preconceived idea. The opportunities were limited almost entirely to printing, shoemaking and carpentry with little consideration for individuality and ability.

Fortunately, however, through experience and revisions in the economic standards it has been recognized that there are many trades in which the deaf are capable of adjusting. One factor directly responsible for this alteration of opinion has been the realization that some trades have gradually become obsolete, and the deaf, unable to utilize profitably their training at these trades, have had to adjust in other vocations. Through this change in circumstances a new field of vocational approach has been unearthed with unlimited possibilities for development.

VOCATIONAL TRAINING FOR DEAF GIRLS

Vocational training for deaf girls presents a problem of considerable complexity and in many respects is more difficult to approach than vocational training for deaf boys. The reason for this lies primarily in the fact that the vocational possibilities for women are decidedly few as compared with those for men, and the competition is therefore more intense. The effect of this competition between deaf and hearing women, with the exception of a select minority, raises an unsurmountable barrier for the deaf girl, as there are no occupations where this competition might be minimized. There are no vocations as such where the deaf girl need not compete with the hearing. This does not hold true for deaf boys where the vocational opportunities are most varied.

Also, many fields of work in which women are capable of adjusting are not dependent upon any systematized vocational preparation. The results of this survey show that in the great majority of cases where deaf girls have received employment, the work is in a factory or is of general routine nature and of varied types. The impracticability as well as the impossibility of organizing a vocational training program to meet these requirements is obvious.

The trades now taught at the Illinois School for the Deaf are: sewing and dressmaking, cooking, millinery, photography, art, and just recently, beauty culture.

It is surprising to note that in the suggestions for new trades to be adopted typing is by far the most popular. In spite of this frequent impression, however, there are many factors which make the adoption of this training inadvisable. In the first place the severe language handicap of the deaf associated with their difficulty with grammatical construction and their restricted vocabulary, would greatly retard individual expression. This disability, said to be a common fault with many hearing girls, is greatly increased in the deaf. In the second place, the possibility of secretarial or general office work where communication is an essential factor, would be immediately eliminated. And thirdly, in the overcrowded field of typing where the advantages are strictly with the hearing girl, the competition would be decidedly prohibitive.

Suggestions were also made for the adoption of a course in beauty culture and this study has already been added to the present curriculum. There do not appear to be any outstanding reasons why a deaf girl should not be able to adjust at this trade and the results of a few years experimentation in conjunction with the practical application of this training will determine the merits of this vocation.

Before any constructive program can be conceived, however, it is necessary first to determine the purpose or aim of vocational training for deaf girls. What is to be accomplished by this training and what measures are to be adopted to assure the derivation of the greatest benefits? There are three major aspects for the consideration of this problem.

Should the purpose of vocational training be to prepare the

deaf girl for some specific trade whereby she may contribute to her financial independence or should the purpose be to train her in those arts of domesticity which are essential to economic and social welfare, such as cooking, dressmaking and millinery? Or, should an attempt be made to effect a combination of both procedures, emphasizing, in addition to those domestic courses, instruction in crocheting, knitting, hooking and weaving? Training in these handiworks might enable a deaf girl to obtain some financial security and at the same time provide an interesting recreational hobby to occupy much leisure time that might otherwise be dull and fruitless.

It is doubtful whether or not an unqualified opinion can be adopted for any one of the three programs just presented since every girl presents an individual problem for consideration and the factors having the most direct influence are so diversified as to prevent general classification, and also since the vocational possibilities are limited to such an extent that competition is a restrictive element.

Photography, however, does not have much commercial value for deaf girls and might well be supplanted by those handiworks previously mentioned. Other than the addition of crocheting, knitting, hooking, and weaving to the present vocational training program, and the exception of photography, there are no major revisions to be offered as a result of this study.

DISCUSSION

Since school offers practically the only opportunity for supervised vocational study and this period is relatively short as compared with other school activities, the degree of industrial success of a deaf child will largely depend upon the training he receives in school. For this reason it seems essential that every effort should be made to determine at an early age so far as possible the natural abilities and tendencies of an individual. These abilities can then be directed and developed so that by the time the child is to leave school he is prepared to the extent of his capacity in the profession or trade for which he is best adapted.

The practice of arbitrarily placing a child in a particular trade or shifting him year after year from one trade to another can

only result in the majority of cases in producing an individual who is totally unprepared to meet the social and industrial obligations of the community in which he resides.

The effort made to determine natural abilities at an early age may rightfully be termed pre-vocational analysis. The consensus of opinion with reference to the age at which a deaf child should begin to receive vocational training is fairly well established by major deaf institutions throughout the country at 13 or 14 years, regardless of grade placement. If the first two years of this training could be devoted to a pre-vocational study of individual abilities, it seems likely that the remaining time would be more advantageously employed. During this period of analysis a child would have an opportunity for securing some information and experience in all available vocations and at the same time a careful record would be kept of his progress at each endeavor. A series of vocational tests given to determine oustanding aptitudes in general occupational classifications might greatly facilitate an ultimate solution. This analysis, conducted under the supervision of a competent vocational guidance committee or counselor, would be of great assistance to the deaf child in planning and choosing his course of study so that by the time he is fifteen or sixteen years of age, the data accumulated during the intervening period could be carefully analyzed and the major objective or vocational subject then determined.

This decision would be materially affected also by a report of the field worker who would be responsible for a presentation of the social and industrial opportunities in the community from which a child came and to which he will most likely return.

The efforts of these combined forces in a pre-vocational study should undoubtedly prove to be far superior to the present system of training and should result in a more complete preparation and ultimate vocational adjustment for the deaf.

The establishment of an employment bureau in connection with the school for the purpose of contacting future employers in an endeavor to assure placement of students upon the completion of their school training, would greatly reduce one of the most serious problems to be confronted by the deaf. Such a bureau might also effect a considerable reduction in the operating budget of a state

institution. There are many deaf students who, unable to meet the academic requirements for graduation, are content to remain under the custodial care of the state until they reach the maximum age limit. The cost of maintaining a deaf child in the Illinois School for the Deaf for one year is in excess of six hundred dollars.

When one realizes also, on the basis of the replies to questionnaires, that only thirty-seven per cent of the students do graduate, the cost of this extended institutional education assumes an appreciable proportion and should be considered.

The educational and vocational achievements of the deaf should be considered separately, as the degree of proficiency in the one is probably not comparable to the degree of proficiency in the other.

The establishment of an employment bureau for the purpose of facilitating the occupational endeavors of the deaf would be of considerable benefit to the students. It would also greatly alleviate the financial burden of the institution in placing those students who have reached the saturation point in their academic progress and are consequently unable to meet the requirements for graduation. When these students attain a sufficient degree of proficiency in a vocation, they should be given the opportunity for assuming industrial responsibility since further educational instruction would not seem warranted.

SUMMARY

The results of the questionnaire sent to former students with regard to their vocational activities reveal that while many are self-supporting, only eight per cent are working at the trade acquired at school. For this reason it seems necessary to suggest a revision in the present system of training to utilize more fully the facilities and opportunities available at the school. These suggestions have been discussed in some detail and need only be summarized.

1. A pre-vocational analysis of all students to determine special abilities.

2. The establishment of a field worker to report on the vocational possibilities in the various communities and to serve

as an intermediary for those students seeking employment.

3. The establishment of a vocational committee or counselor to aid students in their selection of a vocation and to assume responsibility of the employment bureau.

4. The discarding of training in less practical trades. For boys this would involve cabinet making and baking, while for the girls photography would be discontinued.

5. If training in printing is to be continued as a vocational subject, the present equipment should be replaced with more modern devices where necessary in order that this training may have some practical value for a boy who is seeking employment in a modern printing office.

6. The adoption of training in more practical trades. For boys this would involve tailoring, barbering, auto mechanics, and farming. For girls, crocheting, knitting, hooking, and weaving would be taught.

7. The establishment of an employment bureau to contact future employers and to place students on their completion of school training.

PREDICTION OF COLLEGE ACHIEVEMENT FROM DATA COLLECTED DURING THE SECONDARY SCHOOL PERIOD

F. H. FINCH and C. L. NEMZEK

University of Minnesota

DESPITE the fact that intelligence tests have been extensively used for more than a decade, there is a paucity of information based upon direct investigation of the value of tests given early in an individual's educational career for the prognosis of scholastic achievement at a much later period. If results obtained from the administration of intelligence tests during the elementary or secondary school periods of training could be employed for effective prediction of performance at much later periods, as, for example, at the college level, the value of such results for guidance obviously would be very great.

Below are presented the results of an investigation of the effectiveness of prediction of achievement in the freshman year of college from test results obtained at the beginning of the high school careers of the subjects being studied, together with the effectiveness of prediction from data not obtainable until the end of the secondary school period. Records were available for 118 graduates of the University of Minnesota High School who entered the College of Science, Literature, and the Arts of the University of Minnesota. Ninety of these subjects had been in University High School four years, and the remaining 28 had spent three years in University High School. The data consisted of (1) honor point averages based upon freshman marks in the College of Science, Literature, and the Arts of the University of Minnesota, (2) honor point averages based upon high school marks earned at the University of Minnesota High School, (3) University of Minnesota College Aptitude Test percentile ranks obtained at approximately the time of entrance to the University of Min-

nesota, and (4) intelligence quotients obtained at the time of entrance to University High School.

University honor point averages were computed in no case where the total number of quarter hours credit earned by an individual fell below 23 (15 may be considered a normal program for a quarter). This serves to increase the reliability of the measure, but at the same time to restrict the range of talent involved, as any student who did particularly poor work in college would not remain to complete courses leading to the above amount of credit. The high school honor point averages were based on all work completed at the University High School. Sixteen or more credits were included in the computations for each of the 90 graduates who spent four years in the high school, and 12 or more credits for each of the remaining 28 graduates who spent three years in the high school.

High school and university honor point averages were calculated by assigning arbitrarily three points to each quarter hour of A work, two points to each quarter hour of B, one point to each quarter hour of C, zero points to each quarter hour of D, and minus one point to each quarter hour of F. This treatment of marks affords numerical indices with a possible range from minus one to plus three.

The IQ's used were obtained from the administration of five group intelligence tests. The tests used were Army Alpha, Form 8; Haggerty Delta 2; Pressey Senior Classification; Terman Group Test, Form A; and the Miller Mental Ability Test, Form A. The raw IQ's were equated according to Miller's[1] method by reading the values from the table presented by Kefauver.[2] For the present study the middle equated IQ of each pupil was employed.

Of the 118 cases, 90 were administered the five group intelligence tests prior to entrance to the ninth grade, or more than four years before matriculation at the University of Minnesota,

[1] Miller, W. S. "The Variation and Significance of Intelligence Quotients Obtained from Group Tests." *Journal of Educational Psychology,* XV (1924), 359-366.

[2] Kefauver, Grayson N. "Need of Equating Intelligence Quotients Obtained from Group Tests." *Journal of Educational Research,* XIX (1929), 92-101.

and the remaining 28 were tested prior to entrance to the tenth grade, or more than three years before matriculation at the University of Minnesota.

Zero order coefficients of correlation between the six pairs of variables represented by (1) university honor point averages, (2) high school honor point averages, (3) college aptitude test percentile ranks, and (4) IQ's, together with corresponding means and standard deviations, are included in Table I separately for the total group of 118 cases and for the sub-group of 90 cases. The data indicate that honor point averages earned in high school are distinctly superior to college aptitude test percentile ranks or to IQ's for the prediction of honor point averages based upon freshman university grades, and that IQ's obtained from group intelligence tests administered three or four years before entrance to university form at least as good a basis for prognosis as college aptitude test percentile ranks obtained at approximately the time of entrance to university.

The comparative predictive value of high school honor point averages, college aptitude test percentile ranks, and IQ's may be shown more clearly by computing probable errors of estimate or predictive indices. For instance, if the coefficient of correlation between two variables is 1.00, the probable error of estimating an individual's score on X from his score on Y would be .00; on the other hand, if the coefficient of correlation between two variables is .00, the probable error of estimating an individual's score on X from his score on Y would be the probable error of the X distribution. Therefore, the probable error of estimate may be used to determine the comparative prognostic value of various measures when their correlation with the variable to be predicted is known. It is clearly evident that the lower the error of prediction the more desirable is the measure employed. In the present instance in Table II are shown probable errors of estimating freshman university honor point averages from high school honor point averages, college aptitude test percentile ranks, and IQ's. The figures demonstrate that the high school honor point averages have the greatest value for prediction, because the probable errors of estimating university honor point averages from the high school honor point averages are decidedly lower than the

TABLE I

Zero Order Coefficients of Correlation Based Upon (1) University Honor Point Averages, (2) High School Honor Point Averages, (3) College Aptitude Test Percentile Ranks, and (4) IQ's, with Means and Standard Deviations of the Variables

	N=118	N=90
r_{12}	.77 ± .03	.79 ± .04
r_{13}	.42 ± .05	.40 ± .06
r_{14}	.48 ± .05	.48 ± .05
r_{23}	.41 ± .05	.46 ± .06
r_{24}	.46 ± .05	.49 ± .05
r_{34}	.62 ± .04	.62 ± .04
Mean 1	1.26	1.26
Mean 2	1.52	1.56
Mean 3	63.6	60.7
Mean 4	119.3	119.2
Sigma 1	.67	.65
Sigma 2	.57	.56
Sigma 3	27.51	28.16
Sigma 4	10.40	10.53

*Attention is called to the fact that coefficients involving this variable have been computed by the product-moment formula, despite the fact that the scores are expressed as ranks. An error of small magnitude results.

**Throughout this discussion (1) designates university honor point averages; (2), high school honor point averages; (3), college aptitude test percentile ranks; and (4), IQ's.

probable errors of estimate obtained from the college aptitude test percentile ranks or the IQ's.

TABLE II

Predictive Indices and Probable Errors of Estimating University Honor Point Averages from High School Honor Point Averages, College Aptitude Test Percentile Ranks, and IQ's

	N=118			N=90		
	r	P. E. (Est.)	P. I.	r	P. E. (Est.)	P. I.
r_{12}	.77	.29	.36	.79	.27	.39
r_{13}	.42	.41	.09	.40	.40	.08
r_{14}	.48	.40	.12	.48	.38	.12

The same facts may also be portrayed by the use of predictive indices ($1-\sqrt{1-r^2}$) which show that predictions based upon the high school honor point averages are 36 and 39 per cent better than pure guesses for the two groups (N=118: N=90), respectively; whereas, predictions based upon the IQ's are 12 per cent better than pure guess, and the college aptitude test percentile ranks afford predictions that are only 9 and 8 per cent superior to pure guesses. In view of the fact that as the predictive indices increase in magnitude the errors of prediction decrease, these data indicate that the high school honor point averages are superior to the IQ's and the college aptitude test percentile ranks for purposes of prognosis; likewise, the IQ's and the college aptitude test percentile ranks are practically of equal effectiveness for prediction.

While high school honor point averages are decidedly superior to either college aptitude test percentile ranks or IQ's for the prediction of freshman marks in the case of this group of 118 students who entered the College of Science, Literature, and the Arts from the University of Minnesota High School, it is not possible to conclude from the foregoing that prediction may not be improved by employing one or both of the sets of test results together with high school honor point averages, since the various measures may

in part refer to different factors, each of which plays a part in determining college achievement. A combination of the available measures by appropriate statistical methods may be markedly superior for purposes of prediction to any one of the variables considered alone. To investigate this possibility of increasing prediction by combining two or three variables, various multiple R's were calculated (Table III).

TABLE III

Multiple Coefficients of Correlation Based Upon (1) University Honor Point Averages, (2) High School Honor Point Averages, (3) College Aptitude Test Percentile Ranks, and (4) IQ's

	N=118	N=90
$R_{1(23)}$	.779	.791
$R_{1(24)}$	.783	.797
$R_{1(34)}$	.505	.497
$R_{1(234)}$	.784	.797

The multiple coefficients of correlation indicate that the increase in prediction obtained by combining college aptitude test percentile ranks and IQ's with high school honor point averages is of such a small magnitude that it has negligible practical value. In other words, the testing data—college aptitude test percentile ranks and IQ's—do not contribute to a significant degree for predictive purposes any factors not measured or contributed by the high school honor point averages available for the subjects of the present investigation.

When contrasted with the figures usually found in the literature dealing with the relationship between high school and college achievement, the correlations here reported seem surprisingly high.[3] A partial explanation of the fact that a closer relationship

[3] A summary of these studies will be found in Douglass, Harl R. *The Relation of High School Preparation and Certain Other Factors to Academic Success at the University of Oregon.* University of Oregon Publications, Education Series, Vol. III, No. 1, pp. 14-22.

than that reported by most other workers has been observed to exist in this case may be found in the relative uniformity of standards and meaning of marks within University High School. It is quite improbable that a combination of marks from a wide variety of schools will afford a satisfactory means of predicting further achievement until marking systems are decidedly improved. The present results do demonstrate that it is possible under favorable conditions to obtain from marks assigned by high school teachers a predictive measure of distinct value.

The correlations between IQ's and university honor point averages are of particular interest, since they compare quite favorably with the results obtained from the University of Minnesota College Aptitude Test, and with results obtained in previous investigations of the effectiveness of prediction of college achievement from test data, despite the fact that the testing upon which these IQ's are based was done several years before the subjects entered upon their college courses.

THE RELIABILITY OF THE RORSCHACH INK-BLOT TEST[1]

MARGUERITE R. HERTZ

Western Reserve University

The Rorschach Ink-blot Test [2] has been applied in many fields—abnormal psychology, psychiatry, mental tests, psychology of perception, heredity, childhood and adolescence, educational psychology and vocational guidance. The technique, the symptomatic value of the various categories and the interpretation of the test as a whole have been subject to extensive investigation. The method has been found of value in making personality studies of school children and adults, in studying racial differences, adolescent changes in personality, and the influence of environment and heredity on mental make-up, also in diagnosing general intelligence, types of intelligence and grades of mental deficiency. It has likewise been studied in connection with various other typological systems—constitutional types, eidetic types, graphological types and form-color types. It has been used to reveal imagination, inventiveness, fantasy living, poetic talent, and various æsthetic values, and to detect mental and moral defects, schizophrenic, neurotic, psychotic traits and the like. Finally it has been used as a clinical instrument for making diagnosis and has been found of value in psychoanalysis.[3]

A review of the literature discloses a shocking scarcity of data on the reliability of the method. With few exceptions, it has not even been challenged. Despite extensive investigation and appli-

[1] Published through the courtesy of the Richman Fund.

[2] For a description of the method, reference should be made to *Rorschach* (18), *Rorschach and Oberholzer* (19). Summaries of the method may be found in *Behn-Eschenburg* (4), *Enke* (5), *Soukup* (20), *Müller* (12), *Oeser* (14), *Löpfe* (10), *Loosli-Usteri* (9), *Pfahler* (15), *Beck* (2, 3), *Vernon* (22), and *Hertz* (8).

[3] Subsequent investigations of the Rorschach method are summarized by *Beck* (2, 3), *Vernon* (22), and *Hertz* (8).

cation little thought seems to have been given to the problem of whether or not it diagnoses personality consistently and reliably. Without doubt the results of the studies made with the method depend upon its reliability and must be interpreted with caution until that is established. Before the Rorschach test can be used as an instrument of measurement and diagnosis, it must show consistency of results when applied repeatedly to the same individuals under similar conditions.

The problem of reliability

A number of factors make for inconsistency of results. They include 1) lack of standardization of procedure, 2) too limited a sampling of the individual's behavior, 3) unreliability of the sampling of the individual's behavior because of the variability of his performance, and 4) lack of objectivity in scoring.

It was observed in working with the Rorschach Test that there appeared to be great variability in the reactions of divers individuals to the test and in the reactions of the same individual at different sittings. It was thought that such wide and variable performance might not give an accurate index of the individual's habitual behavior or personality. It was of greatest importance, therefore, to study the variability-performance factor to determine whether the sample of the individual's behavior appearing at one sitting was consistent with that at another time.

It was likewise thought that ten pictures might be too small a number of components for a test instrument to elicit a good sampling of behavior. Behavior under the test conditions might not be sufficiently extensive to give a satisfactory and reliable score.

Finally, it was observed from preliminary work with the test that the scoring was highly subjective and might be too dependent upon the "intuition" of the examiner at the expense of its reliability.

All these factors had to be taken into consideration, in studying the reliability of the ink-blot test.

Reliability is generally reported in terms of the coefficient of correlation, i.e., the correlation of the scores of the same individuals upon two successive and similar tests. In the absence of duplicate tests either the same material is used twice or the split test method

is employed in which the test is divided into halves as comparable as possible, the score determined on each half, and those half scores correlated thus giving the reliability of half of the test.

Methods used in other investigations

Two studies are reported in which the Rorschach blots were given twice. *Mira* (11) applied the test to a group of subjects twice a fortnight apart. In certain groups he observed consistency of response. He considered the amount of change indicative of the stability of the individual. He does not include any statistical evidence of these findings. In like manner, *Wertham and Bleuler* (25) gave two applications of the test to determine differences in reactions of individuals under normal circumstances and when under the drug, mescaline. They reported comparatively close agreement in the two sets of responses, with one exception. The slight differences which did occur, did not materially affect the interpretation as a whole. In this study, also, no statistical computations are presented.

Behn-Eschenburg (4) reported the use of a parallel series of ink-blots which gave the same results as the standard Rorschach series. His claim is not, however, statistically substantiated. Other series of blots supposed to be similar to the original Rorschach blots have been used by *Roemer* (16, 17), *Struve* (21), *Gordon and Norman* (7) and *Weil* (23). These have not, however, been standardized against the Rorschach series, and there is no statistical evidence that they are actually similar in nature, equally difficult, and basically measuring the same factors.

The only statistical study to the writer's knowledge which specifically deals with the reliability of the method is that of *Vernon* (22). With 90 subjects, 25 male students of Yale University, 48 male students of Harvard, and 17 male and female adults in England, he used the corrected split-half method and tested the reliability of the Rorschach categories. He divided each subject's scores into two halves of five blots, I, III, V, VI, and X in one series, and II, IV, VII, VIII, and IX, in a second series, correlated them and used the Spearman-Brown correction to show the predicted correlation of the whole test with a second similar test. Scores for his three groups of subjects were tabulated separately.

He obtained the highest correlation for the number of responses (R) +0.91, indicating that if a parallel series were available, the average subject might attain the same total number of responses as on the original set. The reliabilities of the other scores were not as satisfactory, the per cent W (see Table II for symbols) being 0.74, the percentage of M, O, and P falling between 0.70 and 0.60, and the other scores less, indicating that scores might be very different for the two sets. (See Table IV) Vernon ascribes this unsatisfactory reliability of the different categories especially to the subjectivity of the scoring and to the shortness of the test. Because of this latter failing he recommends a parallel set or additional blots. He further states that psychometric standards of reliability do not "properly apply to the test" because Rorschach never intended each separate category to be isolated and treated as a quantitative variable. On the contrary he specifically insisted that no diagnosis should be made without considering the total psychogram. Vernon admits, however, "that if the test is to be regarded as a test at all, or if it claims any objective validity, it must in the future be modified in such a way that the reliabilities of the chief categories of response may achieve a level of at least 0.70—0.80." [4]

In the investigation of the Rorschach Test planned by the Brush Foundation [5] the split-test method was used in the preliminary experiment described in a previous paper. [6] Briefly, the test was administered to 70 subjects constituting a miscellaneous "normal" group, to 30 patients referred to the Psychological Clinic of Western Reserve University, and to 50 patients from the Neuropsychiatric Clinic of Lakeside Hospital, Cleveland, in two sittings, a set of odd-numbered blots being given at one time and after an interval of one to two weeks, the set of even-numbered blots. The

[4] *Vernon*, P. E. The Rorschach Ink-blot Test. British J. Med. Psychol. XIII, Part III, 1933, p. 184.

[5] Research was begun by the writer while a Fellow in Psychology at the Brush Foundation, Western Reserve University, and continued under the direction of Dr. L. Dewey Anderson. The Foundation is now called the Developmental Health Inquiry of the Associated Foundations.

[6] *Hertz*, M. The Standardization of the Procedure and of the Scoring of the Rorschach Ink-blot Test. Western Reserve Library, Cleveland.

answers were scored for those general test factors which required no norms or frequency tabulations, namely for the per cent W, per cent D (any detail answers), percentage of DS, M, C (any color answers), and per cent F (any form answers). The test factors were studied without their qualifications and their differentiations. The corrected coefficients for these factors showed a fairly high reliability in the three groups. Table I reproduces the reliability coefficients for the factors studied.

TABLE I

*Reliability Coefficients for Rorschach Scores with Correction**

	No. of S.	%W	%D	%DS	%M	%C	%F
Miscellaneous	70	.66	.72	.60	.58	.62	.86
correction*		.80	.83	.75	.73	.77	.91
Psychological Clinic	25	.71	.69	.60	.24	.36	.55
correction*		.83	.82	.75	.39	.53	.71
Neuropsychiatric Clinic	50	.71	.61	.65	.55	.88	.81
correction*		.85	.76	.78	.71	.93	.69

* Brown-Spearman Prophecy Formula $r_{12} = \frac{2r}{1 + r}$

It was because the test in this general form showed satisfactory reliability that the experiment proper was set up. It was thought that if the procedure could be made more uniform and the scoring more objective, the reliability of the test could be definitely established.

Standardization of the Ink-blot Test

Reference must be made to the previous paper for details as to standardization of the method of administering the test and of the method of scoring.[7] Suffice to say here that revision was made of certain steps in the procedure to insure standardized conditions. Record blanks, summary sheets and diagrams were rearranged and improved, a trial blot was introduced into the procedure, time of exposure of each blot was limited to two minutes, uniform direc-

[7] ibid

tions were given, and the test situation was kept as uniform as possible. The test was given to 300 students of Patrick Henry Junior High School, Cleveland, 150 boys and 150 girls, ages ranging from 12, 6 to 16, 5. The test records obtained were subjected to statistical analysis and definite quantitative and qualitative criteria were determined for scoring the Rorschach test factors. Responses were tabulated for each ink-blot, the tabulations constituting "Frequency Tables." These tables represented the standard of normality for the specific age group and were used in scoring W, D, Dr, Do, F, O, I and P. The scoring method was found to be reliable as judged by the extent of agreement between the scores of one judge and those of another on 100 records selected at random.

Procedure in determining the reliability of the test

It was decided to use the split-test method again in determining the reliability of the ink-blot test as had been done in the preliminary work. Two applications of the original set were not advisable since factors of familiarity and practice would enter and prejudice results. No parallel series of blots was available. The split-test method was the only method left.

This procedure has its limitations. The assumption is made that halves of the test are approximately equivalent in difficulty, content, and similar respects. When the Rorschach test was divided into a set of five odd-numbered cards (Set A) and another set of five even-numbered cards (Set B), it was observed that the latter contained one colored blot more than in the former. It is to be noted that *Vernon's* sets (22) were not strictly comparable for the same reason. Second, there are several test factors of diagnostic significance which very obviously could not be studied with this method because their existence depends upon the original progression of the series, the *succession* for example. These factors had to be omitted from this study because of the method of approach.

Using the split-test method was justified because it seemed to be the only method available. No test factors which depended upon the progression of the series were included in this treatment. This method further would reveal the influence of the

variability-performance factor, the adequacy of the instrument, and of the standardization of the procedure in reference to the test factors selected for study.

COMPUTATION OF THE COEFFICIENTS OF RELIABILITY

One hundred records were selected at random from the group of three hundred. Scores of each of these were divided into two parts, Set A containing scores of the answers to the odd numbered cards and Set B containing scores of the responses to the even numbered cards. The two sets of scores were summarized. In all there were two hundred half sets of scores.

The two sets of scores were compared for the following factors:

1. Total number of responses
2. Total number of whole responses
3. Percentage of whole answers
4. Number of normal detail answers
5. Percentage of normal detail answers
6. Number of rare detail answers
7. Percentage of rare detail answers
8. Number of oligophrenic details
9. Percentage of oligophrenic details
10. Number of normal details, rescoring and not considering the oligophrenic category
11. Percentage of normal details, rescoring and not considering the oligophrenic category
12. Number of rare details rescoring and omitting consideration of the oligophrenic category
13. Number of white space details
14. Percentage of white space details
15. Number of good forms
16. Percentage of good forms on the basis of the total responses
17. The %F+ factor of Rorschach
18. Number of movement answers
19. Percentage of movement answers
20. Number of chiaroscuro responses
21. Percentage of chiaroscuro responses
22. Number of color answers of any kind

23. Percentage of color answers of any kind
24. Color score factor of Rorschach
25. Number of animal answers
26. Percentage of animal answers
27. Number of human answers
28. Percentage of human answers
29. Number of answers referring to anatomy
30. Percentage of anatomical answers
31. Number of good original answers
32. Percentage of good original answers
33. Number of popular responses
34. Percentage of popular responses
35. Number of forms verbally mentioned, called "items"
36. The "Erlebnistypus" as suggested by each set

In Table II are presented the reliability coefficients of certain of these factors with correction according to the Brown-Spearman formula.

The Rorschach factors appear to be reliable in most cases. Percentage of anatomical, original, and chiaroscuro answers showed the highest reliability (.9). Satisfactory coefficients (.8) were obtained likewise for number of responses, percentage of whole, rare detail, oligophrenic detail, space detail, color, animal and human form answers, and number of items. Percentage of normal detail, good form, and movement answers and the color score obtained coefficients approximating .70. The lowest coefficient (.6) in the group was that for percentage of popular answers. The normal detail and the rare detail factors appeared to have greater reliability when the oligophrenic detail was omitted. The oligophrenic detail, however, showed high reliability in itself.

Comparing *Erlebnistypen* suggested by each half of the test, the percentage of correspondence is 73.

COMPARISON WITH RESULTS IN PRELIMINARY EXPERIMENT

In the preliminary work, reliability coefficients were found for the percentage of whole, detail, space detail, movement and color answers, treated generally without consideration of qualifications. The split-test method was used, each half of the test being given at a separate sitting.

TABLE II

Reliability Coefficients of Rorschach Test Factors

N=100

Sym.	Test Factors	Coeff.	P.E. ±	Correction*
R	Total number of responses	+.812	.014	+.890
%W	Percentage of *whole* answers	+.717	.032	+.835
%D	Percentage of *normal detail* answers	+.600	.043	+.750
%Dr	Percentage of *rare detail* answers	+.752	.029	+.859
%Do	Percentage of *oligophrenic detail* answers	+.686	.040	+.813
%D	Percentage of *normal detail* answers (No Do)	+.678	.039	+.808
%Dr	Percentage of *rare detail* answers (No Do)	+.814	.024	+.897
%DS	Percentage of *space detail* answers	+.767	.030	+.868
%F	Percentage of *good forms* in relation to R	+.677	.039	+.807
%F+	Percentage of *good forms* in relation to F	+.576	.051	+.730
%M	Percentage of *movement* answers	+.594	.043	+.745
%F(C)	Percentage of *chiaroscuro* answers	+.847	.024	+.917
%C	Percentage of *color* answers	+.677	.039	+.810
C	*Color Score*	+.640	.039	+.763
%A	Percentage of *animal* answers	+.708	.035	+.829
%H	Percentage of *human* answers	+.756	.029	+.861
%Anat.	Percentage of *anatomical* forms	+.944	.013	+.971
%O+	Percentage of *original* answers	+.837	.024	+.911
%P	Percentage of *popular* answers	+.495	.051	+.666
Items	Total number of *Items* given	+.761	.017	+.864

Percentage of correspondence between *personality types* 73%

* Brown-Spearman Prophecy Formula $r_{12} = \frac{2\,r}{1 + r}$

The reliability coefficients with the final results here were for the same factors in more refined form. The same procedure was used except that the test was given in one sitting.

The two sets of coefficients are therefore not comparable. However, they are placed side by side in Table III for observation.

The reliability of the whole, space detail and color answers appear to have been increased. The detail factor is not at all comparable of course, since in the final experiment it referred to normal details as statistically determined whereas in the preliminary work it meant any detail answer.

TABLE III

Coefficients of Reliability (as corrected) in Preliminary and Final Experiments

Group	No. of S.	%W	(%D)	%DS	%M	%C
Preliminary						
Normal	70	.80	(.83)*	.75	.73	.77
Final						
Patrick Henry	300	.89	(.80)*	.87	.76	.81

* not comparable

COMPARISON WITH RESULTS OBTAINED BY VERNON (22)

TABLE IV

Vernon's Reliabilities of Scores on the Rorschach Test *N=90*

Test factors	Average for three groups
W%	0.74
F+%	0.33
M%	0.62
Sum C% *	0.34
A%	0.48
O% *	0.60
P%	0.64
Aver.	0.54
R	0.91

* Note—It is to be noted that these factors are not comparable to similar factors in the present investigation. Color score here is on a percentage basis, whereas in this study, only the percentage of color answers given and the sum color score were used. Again, Vernon used the percentage of original answers given while the percentage of good original answers is herein employed.

Table IV reproduces the correlations found by Vernon averaged for his three groups. It is to be noted that his reliabilities are considerably lower than those obtained in this investigation. His average is 0.54 while the coefficients in the present study average 0.829. The number of responses and the percentage of whole answers appear to be reliable in both studies. The percentage

of original answers given irrespective of qualification as used in Vernon's study has a much lower reliability (0.60) than the percentage of good original answers as herein obtained (0.911). The percentage of popular answers obtains unsatisfactory reliability in both studies. It is to be noted that the %F+ in the present study has a much higher reliability (.73) than the same factor in Vernon's study (.33). This is likewise true of the %A which has a reliability of .829 here while it attained only .48 in the other study. Again the color score appears more reliable (.763) than the sum color score per cent of Vernon.

Possibly the standardization of the procedure, especially limitation of the time of exposure, and standardization and increased objectivity of the scoring can account in part for the better results herein obtained.

In determining the reliability of the *Erlebnistypus,* Vernon used the formula $\frac{M - \text{Sum C}}{R}$ as an index of the degree of introversiveness and extratension. He obtained an average split-half reliability (corrected) of 0.55, a result which did not bear out Rorschach's contention that the single test is adequate to diagnose the personality type. According to our results, a correspondence of 73 per cent was obtained between the types found on each half of the test, which tended to confirm Rorschach.

SUMMARY OF THE RELIABILITY OF THE RORSCHACH TEST

1. The coefficients of reliability of the test factors obtained by the corrected split-test method may be considered satisfactory, since they compare favorably with most of the correlations reported in the literature for intelligence tests or tests of mechanical abilities.

2. The ten component pictures of the Rorschach series are sufficient to elicit a fair sampling of an individual's behavior or personality in terms of the above test factors.

3. Personality traits or behavior patterns in terms of the above factors tend to be consistent and follow a stable and unitary pattern. Hence these patterns may be considered as an accurate index to the individual's habitual responses to the test, and the variability performance factor, though it appears important, evidently does not militate against the reliability of the test.

4. Comparison of the indices of reliability in the preliminary experiment where some general test factors were considered, and in the final work where test factors were studied in their several differentiations and where there was refinement of technique, indicates that the reliability of the test was increased after the procedure was standardized and the scoring was made more objective.

INTERCORRELATIONS OF THE RORSCHACH TEST FACTORS

Despite the fact that Rorschach does not introduce any statistical data in his manual, he does indicate certain definite relationships between the various test factors. An examination of these would serve as a further check on the reliability of the method.

Behn-Eschenburg (4), already referred to, compared the average scores of each test factor of his extreme groups to ascertain if Rorschach's relationships obtained. He did not use the method of correlation. He reported positive correspondence between W and M, per cent F+ and FC, also between per cent F+ and Mr; negative correspondence between W and Dr, per cent A and M, Dr and DS, and between per cent F+ and Dr.

Vernon (22) studied the interrelations of the different test factors by the method of correlation and reported statistically unreliable results. He again attempted to account for these unsatisfactory results in terms of Rorschach's assertions that the relationships are to be found only in the "normal person" and are at times broken down or even completely reversed by emotional disturbances. Since the normal person according to Rorschach is the hypothetical person who never can be found (if he was, he would be abnormal) one should not expect to find the schematic relationships enumerated by him.

Table V contains the correlations which obtain among different factors selected for study in the present investigation.[8]

Comparing the whole answer factor with the others, the highest relationship appears between the movement and the color score and the percentage of good original answers. It corresponds least with the percentage of good forms. Rorschach found the

[8] For treatment of the relationship between the Rorschach test factors and intelligence, see *Hertz, M.* Validity of the Rorschach Test: Intellectual Factors. Western Reserve University, Library, Cleveland, O.

direct proportion between the whole answer and the movement which is suggested here.[9] He did not find clear proportions between W and F or between W and C.[10] Present data agree in that

TABLE V

Intercorrelations of the Rorschach Test Factors*

N=300

	IQ	%F+	%O+	Items	%A	Color Score	Whole Ans.	Move. Ans.
IQ		.460	.398	.186	−.108	.209	.241	.259
%F+	.460		.263	−.003	−.007	.072	.108	.173
%O+	.398	.263		.390	−.436	.305	.380	.509
Items	.186	−.003	.390		−.281	.294	.240	.482
%A	−.108	−.007	−.436	−.281		−.311	−.246	−.294
Color Score	.209	.072	.305	.294	−.311		.393	.177
Whole Ans.	.241	.108	.390	.240	−.246	.393		.424
Move. Ans.	.259	.173	.509	.482	−.294	.177	.424	

* P.E. of the coefficients of correlation for r

0.0 — 0.2 ± .0380

0.3 — 0.5 ± .0323

they do not indicate high correspondence between these factors. Experience with the test bears out the fact that W need not necessarily be associated with good form imagery. Subjects frequently gave whole answers without consideration of the exactness of the forms involved. This is pointed out in the study of the feeble-minded group[11] who, of all the groups studied, gave the highest percentage of whole answers with the lowest percentage of good forms. Again, many normal subjects gave few whole answers with a high percentage of good forms. Data here presented suggest that the more whole answers given, the less the stereotypy and the more originality and productivity, which again is in accord with Rorschach's finding.

[9] Rorschach, H. *Psychodiagnostik*. P. 30 and 54.

[10] ibid p. 30

[11] Hertz, M. Validity of the Rorschach Test: Intellectual Factors. Western Reserve University Library.

The movement factor appears to vary directly with the percentage of good original answers and with the number of items. This appears to be in accord with Rorschach's estimate of this factor.

> "The number of M rises with the productivity of the intelligence, with the wealth of associations, and with the ability to complete new associative connections." [12]

"A direct proportion exists between movement and original answers." [13] Further, a low negative correlation is likewise obtained with the percentage of stereotypy, again showing the adverse relationship suggested by Rorschach.

> "Stereotyped and feeble-minded persons have no movement answers." [14]
>
> "The number of movement answers stands in direct proportion . . . to the variability of the responses, hence in reverse proportion to the animal per cent. And the direct proportion between the number of good original answers and the movement answers stand out even more clearly." [15]

The coefficient of correlation obtained between M and per cent F+ (+.173) does not substantiate Rorschach when he reports

> "With normal subjects, a clear proportion exists between the number of M and the keenness of the form imagery." [16]

Rorschach adds, however, that this direct proportion may be upset wherever emotional disturbances enter.[17] In like manner, Rorschach is not confirmed when he indicates a direct proportion between movement and color factors, but again he indicates exceptions. [18]

Rorschach indicated a proportion between per cent F+ and C, [19] but practically no relationship is indicated by present data. Again,

---

[12] Hertz, M. Validity of the Rorschach Test: Intellectual Factors. Western Reserve University Library. P. 15.

[13] ibid p. 80

[14] ibid p. 15

[15] ibid p. 55

[16] ibid p. 17

[17] ibid p. 17 and 54

[18] ibid p. 21

[19] ibid p. 21

Rorschach finds this correspondence may be reversed by normal subjects with nervous or artistic leanings.[20] The inverse relationship between color and stereotypy indicated by Rorschach is confirmed by the correlation of—.311 which was obtained.

> "Only the stabilized show no color . . . the less color, the more stabilized and the more stereotyped the individual."[21]

Present data also indicate a positive relationship of +.305 with the per cent 0+ which bears this out.

The percentage of stereotypy appears negatively related to per cent 0+, to the number of items, the color score, whole and movement answers. The highest relationships obtained are with the per cent 0+ (—.436) and the color score (—.311). These facts are all suggested by Rorschach. He indicates that artistic and imaginative people have a low per cent A, while the stereotyped and the feeble-minded have a high per cent A.[22]

It should be noted that the number of items shows highest correlation with the per cent 0+ and the movement answers. This factor was introduced into the present study and was not suggested by Rorschach.

SUMMARY FOR THE INTERCORRELATIONS OF THE RORSCHACH TEST FACTORS

1. Intercorrelations range from —.436 to +.509.
2. The whole answer factor shows highest correspondence with movement, color score, and percentage of good original answers.
3. The movement factor is best related to the good original answers and corresponds likewise to items and number of whole answers.
4. The color score is positively related to per cent 0+ and W and negatively related to per cent A.
5. The per cent A shows a negative relationship to all the other factors especially with the per cent 0+, the C and the W.

---

[20] ibid p. 21

[21] ibid p. 22

[22] ibid p. 36, Table VI.

6. The items factor is best related to movement and originality.

7. While definite conclusions cannot be made, the figures suggest many of the findings which Rorschach reported.

GENERAL CONCLUSION

Statistical treatment of the results obtained with the Rorschach test in its modified and standardized form shows the test to be a reliable instrument.

BIBLIOGRAPHY

1. Beck, S. J. Personality Diagnosis by Means of the Rorschach Test. *Amer. J. Orthopsychiat.*, 1930, 1, 81-88.

2. _____., The Rorschach Test and Personality Diagnosis. I. Feeble-minded. *Amer. J. Psychiat.*, 1930, 10, 19-52. Also in *Institute for Child Guidance Studies* (L. G. Lowrey), pp. 222-261. N. Y.: Commonwealth Fund, 1931.

3. _____, The Rorschach Test as Applied to a Feeble-Minded group. *Archives of Psychol.*, 136, 1932, p. 84.

4. Behn-Eschenburg, H., Psychische Schüleruntersuchungen mit dem Formdeutversuch. Ernst, Bircher Verlag, Bern u. Leipsig, 1931, p. 69. Also Huber, Bern.

5. Enke, W. Die Konstitutionstypen im Rorschachschen Experiment. *Zsch. f. d. Neur. u. Psychiat.*, 1927, 108, 645-674.

6. _____, Die Bedeutung des Rorschachschen Formdeutversuches für die Psychotherapie, *Sitzungsber. d. Ges. z. Beförderg. d. ges. Naturw. z. Marburg.* 1927, 62, 621-633, Berlin: O. Elsner Verlag.

7. Gordon, R. G. and Norman, R. M., Some Psychological Experiments on Mental Defectives in Relation to the Perceptual Configurations Which May Underlie Speech. Part II. *Brit. J. Psychol.*, 1932, 23, 85-113.

8. Hertz, M. R., The Rorschach Ink-blot Test. Unpublished Thesis, Western Reserve University Library, Cleveland, Ohio.

9. Loosli-Usteri, M., Le test de Rorschach, appliqué à différents groupes d'enfants de 10-13 ans. *Archives de Psychol.*, 85, 1929, 51-106.

10. Löpfe, A., Uber Rorschachsche Formdeutversuche mit 10-13 jährigen Knaben. *Zsch. f. angew. Psychol.*, 1925, 26, 202-253.

11. Mira, L. E., Sobre el valor del Psicodiagnostico de Rorschach. *Progressos de la Clinica*, 1925, 808-845.

12. Müller, M., Der Rorschachsche Formdeutversuch, seine Schwierigkeiten und Ergebnisse. *Zsch. f. d. ges. Neur und Psychiat.*, 1929, 118, 598-620.

13. Munz, E., Die Reaktion des Pyknickes im Rorschachschen psychodiagnostischen Versuch., *Zsch. f. d. ges. Neur. und Psychiat.*, 1924, 91, 26-92.

14. Oeser, O. A., Some Experiments on the Abstraction of Form and Color. Pt. II. Rorschach Tests. *Brit. J. Psychol.*, 1932, 22, 287-323.

15. Pfahler, G., System der Typenlehren. Grundlegung einer pädagogischen Typenlehre, *Zsch. f. Psychol.*, 1929, Erg. Bd. 15, p. 334.

16. Roemer, G. A., Die Innenwelt einer Persönlichkeit und das Problem ihrer wissenschaftlichen Erschliessung, *Psychol. Rundschau*, 1930, 2, 4-12. *Psychol. Rundschau*, 1930, 2, 33-41. Ibid., 69-76; 101-109. Reprint Die wissenschaftliche Erschliessung der Innenwelt einer Persönlichkeit, Basle: Birkhäuser, 1931, p. 42.

17. ———————, Psychographische Tiefenanalyse einer Grossindustriellen und seiner Stabes. *Prakt. Psychol.*, 1922-3, IV, 16-28.

18. Rorschach, H., Psychodiagnostik. Methodik und Ergebnisse eines wahrnehmungsdiagnostische Experiments. Bern: Bircher, 1921, p. 42. Also enlarged, Bern: Huber, 1932, p. 227.

19. Rorschach, H. and Oberholzer, E., Zur Auswertung des Formdeutversuche für die Psychoanalyse, *Zsch. f. d. ges. Neur. und Psychiat.*, 1923, 82, 240-273. The same, translated: The Application of the Interpretation of Form to Psychoanalysis. *J. Nerv. and Ment. Dis.* 1924, 60, 225-248; 359-379.

20. Soukup, F., The Study of Personality and Rorschach's Test. In Czech. *Čas. lék. česk.*, 1931, 1. 881-887. Abstract: *Zentbl. f. d. ges. Neur.*, 1932, 61, 564.

21. Struve, K., Typische Ablaufsformen des Deutens bei 14- bis 15- jährigen Schulkindern, *Zsch. f. angew. Psychol.*, 1930, 37, 204-274.

22. Vernon, P. E., The Rorschach Ink-blot Test. *Brit. J. Med. Psychol.*, 13, 1933, 89-114; 179-200, 271-291.

23. Weil, H., Wahrnehmungsversuche an Integrierten und Nichtintegrierten. *Zsch. f. Psychol.*, 1929, 111, 1-50.

24. Wells, F. L., The Systematic Description of Personality. Conference on Individual Differences in the Character and Rate of Psychological Development. Washington: National Research Council 1931. p. 52-70.

25. Wertham, F. and Bleuler, M., Inconstancy of the Formal Structure of the Personality: experimental study of the influence of mescaline on the Rorschach test. *Arch. Neur. und psychiat.*, 1932, 28, 52-70.

INTROVERSION-EXTROVERSION AND THE FUNCTIONAL PSYCHOSES

JAMES PAGE

Kings Park State Hospital[1]

HAVING succeeded in reducing as complicated a factor as intelligence to a mere number, psychologists are now attempting to define personality in mathematical terms. Numerous and varied attempts have been made to define and measure the introversion-extroversion aspect of personality.

The question of definition has been discussed by Conklin (1) and Guilford and Braly (2). The most frequently used measuring devices are the questionnaires derived from Freyd's original list (3) of introvert-extrovert traits. These questionnaires stress reactions to social situations. Rorschach (4) has used reactions to ink-blots as a criterion of introversion-extroversion, while Conklin (5) has utilized choice of interests as criteria. Physiological correlates for introversion-extroversion have been offered by Hunt (6) and Kretschmer (7). According to the latter, the asthenic constitution forms the foundation for the introverted or schizoid personality and the pyknic constitution forms the basis for the extroverted or cycloid personality.

In psychiatry, the introversion-extroversion concept has been particularly widespread. Although only partly confirmed by experimental studies (8), the convenient theory—that schizophrenics are extreme introverts and manic-depressives extreme extroverts—is quite prevalent. Diagnosis often depends upon whether the patient is an introvert or an extrovert. Neymann and Kohlstedt (9) have gone as far as to construct a test of introversion-extroversion on the basis of this unsubstantiated theory. Favorable results have been obtained by Gilliland and Morgan (10) with a modification of this test. Campbell (11)

[1] The author wishes to express his appreciation to the medical staff for their excellent cooperation.

using the Heidbreder scale and De Angelis (12) employing the Bernreuter B3-I test, have also found manic-depressive patients to be more extroverted than schizophrenic patients.

These psychological studies are in complete conformity with usual psychiatric findings (8). There is no doubt that schizophrenic patients are more introverted than manic-depressive patients. However, it is open to question whether these two groups, as is frequently held, portray the extremes of introversion-extroversion with normal individuals assuming an intermediate rôle. The present study is concerned primarily with this problem. In addition, we have investigated the following problems by means of standardized tests of introversion-extroversion.

1. The relative occurrence of introversion-extroversion traits in manic-depressive, schizophrenic and normal individuals.
2. Introversion-extroversion ratings and the functional psychoses.
3. Introversion-extroversion and prognosis in psychotic patients.
4. Introversion-extroversion and bodily types.
5. The relative occurrence of specific introvert traits in manic-depressive, schizophrenic and normal individuals.

THE EXPERIMENTAL PROCEDURE

The Heidbreder and the Neymann-Kohlstedt tests of introversion-extroversion were given to 50 manic-depressive patients, 50 schizophrenic patients and 50 hospital attendants. On the Heidbreder scale, the question referring to effeminacy was omitted. In order to facilitate comparison between the two tests, the scoring on the Neymann-Kohlstedt test was slightly modified. According to the authors (9) the score on this test is the algebraic sum of plus and minus scores with extrovert answers being considered plus. We reversed the signs so that the score was the algebraic sum of plus and minus scores with introvert traits being considered plus.

All of the subjects were males and were connected with the Kings Park State Hospital as inmates or employees. The psychotic patients were tested individually by the examiner and the attendants filled out the tests by themselves. Only cooperative

patients with whom satisfactory rapport could be established were tested. The three groups were quite comparable as regards age, education and social status.

2. The attendants in charge of the patients rated each patient on a five point introversion-extroversion scale ranging from (1) extreme extroversion to (5) extreme introversion. Whenever possible, two independent ratings were obtained on each patient.

3. The psychiatrists in charge of each service rated their patients as to bodily build, following Kretschmer's classification. They were instructed to place all the patients that were not clearly of a pyknic or asthenic build in a general mixed group.

QUALITATIVE RESULTS

The various psychotic types reacted quite differently to the testing. The manic-depressive-manic patients tended to be hyperactive, facetious and easily accessible. Responses were given promptly, decisively and in a loud voice. The manic-depressive-depressed patients were also open and accessible but were "slowed-up."

The schizophrenic patients were more suspicious and guarded in their responses. Indecision and perseveration were frequently encountered. Responses were given in a low voice. Reaction time was retarded.

To make sure that our patients were cooperating, we frequently asked them to repeat the question just read and occasionally asked the same question at two different times to determine the reliability of their responses. This check necessitated the elimination of all hebephrenic patients whose responses, on the whole, were entirely unreliable. Our schizophrenic group is consequently composed of only simple, catatonic and paranoid schizophrenic patients.

QUANTITATIVE RESULTS

The average differences between our groups as to introversion-extroversion are given in Table I and the reliability of such differences are reported in Table II. Our results seriously challenge the theory that assumes that manic-depressives are extreme ex-

troverts and schizophrenics extreme introverts with normal individuals assuming an intermediate position. On both our tests the normal group is the most extroverted and the schizophrenic group the most introverted with the manic-depressive group assuming an intermediate rôle. On neither test are the manic-depressive patients reliably more extroverted than the schizophrenic group. The only significant differences are found between the normal and psychotic groups.

In this connection, we may refer to Bernreuter's interesting study (13) which clearly indicates a very high correlation existing between neurotic traits and introversion traits. That the two terms are interchangeable has been brought out by Bernreuter (14) and Stagner (15). Both investigators have reported somewhat higher

TABLE I

*Inter-Comparisons as to Mean Scores**

	Normal N=50	Manic-Depressive N=50	Schizophrenic N=50
Neymann-Kohlstedt			
Mean	−1.15	−.50	3.50
Sigma	9.80	9.00	10.05
Heidbreder			
Mean	−20.30	−11.20	−9.20
Sigma	11.90	14.20	14.50

* On both tests, plus scores indicate introversion.

TABLE II

Reliability of Differences in Mean Scores

	D/sigma diff.	Chances in 100
Neymann-Kohlstedt		
Normal and Manic-Depressive	.35	64
Normal and Schizophrenic	2.35	99
Schizo. and Manic-Depressive	2.11	98
Heidbreder		
Normal and Manic-Depressive	3.49	100
Normal and Schizophrenic	4.19	100
Schizo. and Manic-Depressive	.70	76

correlations between introversion-extroversion tests and neurotic tendency tests, and between different introversion-extroversion tests. Since both psychotic groups have more neurotic traits than normal individuals, we should naturally expect both of them to be more introverted than the normal group.

Our results clearly indicate that of the two psychotic groups, the personality make-up of the manic-depressive group is considerably more similar to the normal personality. Similar findings have been reported by Bond and Abbott (16), Bowman and Raymond (17) and Bond (18).

As regards the marked extroversion of our normal group, we may mention that other investigators (10), (14), (15), (19) and (20) have also found a predominance of extrovert traits in normal individuals. Mental balance tends to be more closely associated with extroversion than with introversion.

ANALYSIS OF TESTS USED

Our results on the Neymann-Kohlstedt test strongly contradict the authors' claim for this test as a diagnostic instrument and question the validity of their criterion of introversion-extroversion. (The test is based on the assumption that all schizophrenes are introverts and all manic-depressives are extroverts.) Only 60 per cent of our schizophrenic group turned in introvert scores and but 44 per cent of our manic-depressive patients developed extrovert scores. Fifty-seven per cent of our normal group reported extrovert scores.

The Heidbreder test differentiated much more clearly between the normal and psychotic groups. In all groups, there was a piling of scores toward the extrovert end, suggesting that the ambivert (zero) score should be shifted toward the extrovert end. On this test, 30 per cent of the schizophrenic patients reported introvert scores, 74 per cent of the manic-depressive patients reported extrovert scores and 98 per cent of the normal individuals reported extrovert scores.

INTROVERSION-EXTROVERSION RATINGS

A common criticism of personality tests of the type employed in this study is that they interpret personality as a mere summa-

tion of traits, whereas personality is more likely a dynamic integration of traits. To test this criticism, we obtained introversion-extroversion ratings on 86 of the 100 patients tested. Since the patients were rated on a five point scale, the ratings were necessarily based on the general personality as a whole rather than on the algebraic summation of traits. This gestalt approach confirms our test findings. As shown in Table III, there is a tendency for schizophrenic patients to be more introverted than manic-depressive patients but the difference is not completely reliable statistically. Of the 42 schizophrenic patients, 29 were rated as introverts, 5 as extroverts and 8 as ambiverts (rating of 3). Nineteen of the 44 manic-depressive patients were rated as introverts, 11 as extroverts and 14 as ambiverts.

TABLE III

Introversion-Extroversion Ratings and Psychoses

	Mean	Sigma	D/sigma diff.	Chances in 100
Schizophrenics	3.60	.86		
Manic-Depressives	3.07	.94	2.72	99

INTROVERSION-EXTROVERSION AND PROGNOSIS

In psychiatry, the feeling is quite prevalent that the more introverted a patient is, the poorer the prognosis and the more extroverted a patient is, the better the prognosis. We have obtained some statistical data (Table IV) on this point by comparing the ratings and test scores of patients who left the hospital within four months after the testing with the ratings and test scores of patients who were still in the hospital.

As is apparent, there is absolutely no relationship between introversion-extroversion and prognosis. In a study of prognosis in manic-depressive psychoses, Steen (21) likewise found no relationship between the shut-in (introverted) personality and recovery. In some earlier unpublished work with psychotic patients, the author found zero correlation coefficients existing between the number of schizoid traits present and psychiatrists' ratings as to recovery.

TABLE IV

Relation of Introversion-Extroversion to Prognosis

	N	Mean Scores Heidbreder	Neymann-Kohlstedt	Ratings
In Hospital	65	−10.00	1.50	3.30
Out of Hospital	35	−11.35	1.95	3.35

INTROVERSION-EXTROVERSION AND BODILY TYPES

Twelve of the patients were rated by their physicians as typical pyknics and 12 as typical asthenics. All of the other patients were either of a mixed type or could not be accurately rated for some reason or other. As brought out in Table V, there is a definite tendency, both as regards ratings and test scores, for asthenic patients to be more introverted than pyknic patients.

TABLE V

Relation of Bodily Build to Introversion-Extroversion

Mean	Pyknic	Asthenic
Ratings	3.25	3.42
Neymann-Kohlstedt	4.33	5.83
Heidbreder	−9.83	−6.66

INTER-GROUP COMPARISONS BASED ON ITEM ANALYSIS

We next determined the relative percentage of each group giving an introvert answer to each specific trait on the two tests. Our three groups were then compared by:

1. Finding the correlation coefficients between the three groups.
2. Determining which traits were reported more frequently by one group as compared with the other two groups.

From the high correlation coefficients reported in Table VI, three conclusions may be drawn:

1. Traits reported frequently or rarely by one group tend to be reported frequently or rarely, respectively, by the other two groups.
2. The similarity between the two psychotic groups is con-

siderably more marked than the similarity between the normal group and either psychotic group.

3. Of the two psychotic groups, the manic-depressive group bears greater resemblance to the normal group.

TABLE VI

Inter-Correlations Based on Item Analysis

	Heidbreder	Neymann-Kohlstedt
Normal and Schizophrenic	.86	.68
Normal and Manic-Depressive	.89	.73
Schizo. and Manic-Depressive	.93	.90

In determining whether certain traits were more characteristic of one group than another, we used the usual criterion, D/sigma difference=at least 3. Our work was considerably simplified by the tables of Edgerton and Paterson (22).

The close relationship existing between the two psychotic groups is again evident. The schizophrenic group failed to report a single introvert trait reliably more frequently than the manic-depressive group on either test.

On the Neymann-Kohlstedt test, the normal group reported one trait (11) (Table VII) reliably more frequently than the manic-depressive group and one other introvert trait (39) more frequently than the schizophrenic group. Traits (1), (38) and (43) were more characteristic of both psychotic groups as compared with the normal group. In addition the manic-depressive group reported traits (12) and (35) and the schizophrenic group reported trait (48) reliably more frequently than the normal group.

TABLE VII

Differentiating Traits on the Neymann-Kohlstedt Test
(Introvert answer designated)

11. No Accept suggestions rather than working them out for yourself.
39. No Act on the spur of the moment.
1. Yes Be by yourself a great deal.
38. Yes Keep quiet when out in company.

43. Yes Avoid trouble rather than face it.
12. Yes Quiet rather than exciting amusements.
35. Yes Read about rather than do a thing.
48. Yes Spend your vacation at some quiet place rather than at a lively resort.

On the Heidbreder test (3) the normal group failed to report a single introvert trait reliably more frequently than the psychotic group. The following 8 introvert traits (Table VIII) were reported reliably more frequently by both psychotic groups as compared with the normal group.

TABLE VIII

Differentiating Traits on Heidbreder Test
("Yes" answers indicate introversion)

7. Introspect, turn your attention inward.
21. Shrink when facing a crisis.
28. Indulge in self pity when things go wrong.
33. Get rattled easily, lose your head in excitement or moments of stress.
43. Pay serious attention to rumors.
44. Believe in mind cures, accept an idealistic philosophy.
49. Are slow in movement.
50. Prefer to read a thing rather than experience it.

In addition the manic-depressive group (Table IX) reported trait (15) and the schizophrenic group reported traits (2), (6), (20), (45) and (51) reliably more frequently than the normal group.

TABLE IX

Other Differentiating Traits on the Heidbreder Test
("Yes" answers indicate introversion)

15. Worry over possible misfortunes.
2. Avoid all occasions for talking before crowds, find it difficult to express yourself in public.
20. Are reticent and retiring, do not talk spontaneously.
45. Talk to yourself.
51. Are suspicious of the motives of others.
6. Hesitate in making decisions on ordinary questions that arise in the course of the day.

In summarizing our analysis of the 103 introvert traits studied, we find:

1. All three groups report 81 or 79 per cent of the traits with about equal frequency.

2. The schizophrenic group reports 17 traits and the manic-depressive group reports 14 introvert traits reliably more frequently than the normal group. Eleven of the traits reported more frequently by the schizophrenic group are also reported more frequently by the manic-depressive group.

3. The normal group reports 2 introvert traits reliably more frequently than either psychotic group.

4. The schizophrenic group fails to report a single introvert trait reliably more frequently than the manic-depressive group.

SUMMARY

The Heidbreder and the Neymann-Kohlstedt tests of introversion-extroversion were given to manic-depressive patients, schizophrenic patients and normal individuals. In addition the psychotic patients were rated as to introversion-extroversion and bodily type. Our results indicate that:

1. Normal individuals are considerably more extroverted than manic-depressive patients who in turn are somewhat more extroverted than schizophrenic patients.

2. Schizophrenic and manic-depressive patients resemble each other considerably more than either resemble normal individuals.

3. The similarity between manic-depressive patients and normal individuals is more marked than the similarity between schizophrenic patients and normal individuals.

4. No relationship exists between introversion-extroversion and prognosis in the functional psychoses.

5. Asthenic patients tend to be more introverted than pyknic patients.

BIBLIOGRAPHY

1. Conklin, E. S.: The Definition of Introversion-Extroversion, and Allied Concepts. J. Abn., and Soc. Psychol., 1923, 17, 367-382.

2. Guilford, J. P., and Braly, K. W.: Extroversion and Introversion. Psychol. Bull., 1930, 27, 96-107.

3. Freyd, M.: Introverts and Extroverts. Psychol. Rev., 1924, 31, 74-87.

4. Beck, S. J.: The Rorschach Method and Personality Organization. Amer. J. Psychiat., 13, 3, 519-532.

5. Conklin, E. S.: The Determination of Normal Extrovert-Introvert Interest Differences. Ped. Sem., 34, 1927, 28-37.

6. Hunt, J. R.: Erethizophrenia and Kolyphrenia. Amer. J. Psychiat., 12, 1932, 475-492.

7. Kretschmer, E.: Physique and Character. Harcourt, Brace and Co., 1925.

8. Page, J., Landis, C., and Katz, S.: Schizophrenic Traits in the Functional Psychoses and in Normal Individuals. Amer. J. Psychiat., 1934 (May). 1213-1225.

9. Neymann, C. A., and Kohlstedt, J. D.: A New Diagnostic Test for Introversion-Extroversion. J. Abn., and Soc. Psychol., 1929, 23, 482-487.

10. Gilliland, A. R., and Morgan, J. J. B.: An Objective Measure of Introversion-Extroversion. J. Abn., and Soc. Psychol., 1932, 26, 296-303.

11. Campbell, K.: The Application of Introversion-Extroversion Tests to the Insane, J. Abn., and Soc. Psychol., 1929, 23, 479-482.

12. De Angelis, L.: M. A. Essay, Columbia University Library, 1932.

13. Bernreuter, R. G.: The Theory, and Construction of the Personality Inventory. J. Soc. Psychol., Nov. 1933, IV, 4, pg. 401.

14. Bernreuter, R. G.: The Imbrication of Tests of Introversion-Extroversion and Neurotic Tendency. To appear soon in the J. Soc. Psychol.

15. Stagner, R.: The Inter-Correlations of Some Standardized Personality Tests. J. Appl. Psychol., 1932, 16, 453-464.

16. Bond, E. D., and Abbot, E. S.: A Comparison of Personality Characteristics in Dementia Præcox and Manic-Depressive Psychoses. Am. J. Insan., 1911, 6, 359-371.

17. Bowman, K. M., and Raymond, A. F.: A Statistical Study of the Personality in Schizophrenic Patients. Schizophrenia, Proc. Assoc. Research Ner., and Ment. Diseases, 1931, 10, 48-74. Williams and Wilkins, Baltimore.

18. Bond, E. D.: Personality and Outcome in Two Hundred Consecutive Cases. Am. J. Insanity, 1912-13, 69, 731-738.

19. Heidbreder, E.: Measuring Introversion and Extroversion. J. Abn. and Soc. Psychol., 1926, 21, 120-134.

20. Root, A. R., and Root, E. B.: A Study of the Neymann-Kohlstedt Diagnostic Test for Introversion-Extroversion. J. Abn. and Soc. Psychol., 1931, 26, 415-421.

21. Steen, R. R.: Prognosis in Manic-Depressive Psychoses. Psychiat. Quart., 1933, 7, 419-429.

22. Edgerton, H. A., and Paterson, D. G.: Table of Standard Errors, and Probable Errors of Percentages for Varying Numbers of Cases. J. Appl. Psychol., 1926, 10, 378-391.

RECOGNITION OF MELODIES AND TONAL MEMORY

EDNA WILLIS McELWEE

New York City

ONE of the purposes of a music club, which the writer conducted for ninth year students in a Junior High school at Canton, Ohio, was to familiarize the members with classical music. In that connection a test for recognition of melodies was devised. The thirty members of the music club may be considered a selected group of children for all of them took music lessons in either piano, violin, cornet or voice.

At every meeting three or four new melodies were played and the name of each of them given. The part of the selection chosen was limited to the first melodic sentence, which varied from four to ten bars. After the preliminary instruction in the new melodies, several melodies were given, the new ones being interspersed among familiar ones. The children wrote the name of each melody as it was played.

This preliminary step was repeated several times until a graded series of ten new and ten familiar melodies had been obtained. Following are the melodies included in the final test. The table also gives the per cent of the group which recognized each melody.

The test for recognition of melodies was now given to the 150 pupils in the seventh and eighth grades of the same Junior High school. During the preliminary instruction the new melodies were presented in a different order from that in which they appeared in the test in order that the pupil with a good memory for names would have no advantage.

On the same day the test for tonal memory, one of the Seashore tests for musical traits, was given to the seventh and eighth grade pupils. In this test a series of two, three, four or five notes is played and then repeated with one of the notes being changed.

TABLE I

Test for Recognition of Melodies

Melody	Selection	Per Cent Correct
1 N	Cuckoo Song	100
2 F	America	95
3 N	Funeral March	90
4 N	Cradle Song	85
5 F	Dixie	80
6 N	Spring Song	75
7 F	Perfect Day	70
8 N	Dancing Doll	65
9 F	Holy Holy Holy	60
10 N	Chiming Bell	55
11 N	Boat Song	50
12 F	Sweet and Low	45
13 N	To My Beloved	40
14 N	Angel's Serenade	35
15 F	Lead Kindly Light	30
16 N	Narcissus	25
17 F	The Land of the Sky Blue Water	20
18 F	Nellie Gray	15
19 F	I Love You Truly	10
20 F	The Last Rose of Summer	5
	N—new melody	
	F—familiar melody	

The listener records by number which note has been changed in the second playing.

Percentile scores were then calculated for each pupil on both the test for recognition of melodies and for tonal memory. Using the product-moment formula a correlation of $.62 \pm .03$ was found. These results show that the ability to recognize melodies by name has a definite relation to other musical traits, and in particular to that of tonal memory.

NOTE ON THE STANDARD ERROR OF THE DIFFERENCE BETWEEN COEFFICIENTS OF VARIATION OF CORRELATED VARIABLES

JOSEPH ZUBIN

Psychiatric Institute, Columbia University

IN MANY experiments in Education and Psychology the crucial comparison is made between variabilities rather than between central tendencies. One of the measures of variability in common use, especially when the central tendencies of the compared variables are unequal, is the coefficient of variation, $V_\nu = \sigma/M$. The standard error of this coefficient is given by Yule [1] but the standard error of the difference between the coefficients of two correlated variables has not been developed yet.[2] The purpose of this note is to derive the equation for this standard error.

The general expression for the standard error of two correlated variables is

$$\sigma^2_{(\nu_1-\nu_2)} = \sigma^2_{\nu_1} + \sigma^2_{\nu_2} - 2r_{\nu_1\nu_2}\sigma_{\nu_1}\sigma_{\nu_2} \tag{1}$$

Utilizing the equations derived by Pearson[3] under the single assumption that quantities involving $1/N$ are negligible in comparison with values involving $1/\sqrt{N}$,

(2) $\sigma^2_{\nu_1} = \nu_1^2/N \cdot \{(\beta_2 - 1)/4 + \nu_1^2 - \nu_1\sqrt{\beta_1}\}$ and a similar expression can be written for $\sigma^2_{\nu_2}$.

Now $r_{\nu_1\nu_2}\sigma_{\nu_1}\sigma_{\nu_2} = E\left\{\delta\left[\frac{\sigma_1}{M_1}\right]\delta\left[\frac{\sigma_2}{M_2}\right]\right\}$ where E represents

---

[1] Yule, G. U., An Introduction to the Theory of Statistics, 1927 Edition, p. 351.

[2] The equation to be derived has appeared without proof in—Some Effects of Incentives, A Study of Individual Differences in Rivalry, p. 33, Teachers College Contributions to Education no. 532., 1932, by the present writer.

[3] Pearson, K. On the probable errors of frequency constants, Biometrika vol. 9, p. 14.

the sum of all cross product values in all samples, divided by the number of samples. Utilizing Pearson's equations again, we obtain

$$(3)\quad r_{\nu_1\nu_2}\sigma_{\nu_1}\sigma_{\nu_2}=\nu_1\nu_2\{[P_{22}/P_{02}P_{20}-1]/4-r_{12}\nu_1\nu_2-P_{12}/M_1 P_{02}-P_{21}/M_2P_{20}\}$$

Substituting in (1)

$$(4)\quad \sigma^2_{(\nu_1-\nu_2)}=1/N\cdot\{\nu_1^2[(\beta_2-1)/4+\nu_1^2-\nu_1\sqrt{\beta_1}]+\nu_2^2[(\beta_2'-1)/4+\nu_2^2-\nu_2\sqrt{\beta_1'}]-2\nu_1\nu_2[(P_{22}/P_{20}P_{02}-1)/4+r_{12}\nu_1\nu_2-P_{12}/M_1\sigma_2^2-P_{21}/M_2\sigma_1^2]\}$$

For linear and homoscedastic regression,

$$P_{21}=r_{12}\sigma_x^2\sigma_y\sqrt{\beta_1};$$
$$P_{12}=r_{12}\sigma_x\sigma_y^2\sqrt{\beta_1'} \quad \text{and} \quad \beta_2=\beta_2'$$
$$P_{22}/P_{20}P_{02}=r_{12}^2(\beta_2-1)+1$$

Substituting

$$(5)\quad \sigma^2_{(\nu_1-\nu_2)}=1/N\cdot\{\nu_1^2[(\beta_2-1)/4+\nu_1^2-\nu_1\sqrt{\beta_1}]+\nu_2^2[(\beta_2'-1)/4+\nu_2^2-\nu_2\sqrt{\beta_1'}]-2r_{12}\nu_1\nu_2[r_{12}(\beta_2-1)/4+\nu_1\nu_2-\nu_1\sqrt{\beta_1'}-\nu_2\sqrt{\beta_1}]\}$$

For symmetrical distributions, $\beta_1=\beta_1'=0$.

$$(6)\quad \sigma^2_{(\nu_1-\nu_2)}=1/N\cdot\{\nu_1^2[(\beta_2-1)/4+\nu_1^2]+\nu_2^2[(\beta_2'-1)/4+\nu_2^2-2r_{12}\nu_1\nu_2[r_{12}(\beta_2-1)/4+\nu_1\nu_2]\}$$

and for normal distributions

$$(7)\quad \sigma^2_{(\nu_1-\nu_2)}=1/2N\cdot\{\nu_1^2[1+2\nu_1^2]+\nu_2^2[1+2\nu_2^2]-2r_{12}\nu_1\nu_2[r_{12}+2\nu_1\nu_2]\}$$

Following the usual convention of taking $\nu_1=100\sigma_1/M$, instead of σ_1/M_1 as we have done,

$$(8)\quad \sigma^2_{(\nu_1-\nu_2)}=\frac{1}{2N}\cdot\left\{\nu_1^2\left[1+\frac{2\nu_1^2}{(100)^2}\right]+\nu_2^2\left[1+2\frac{\nu_2^2}{(100)^2}\right]-2r_{12}\nu_1\nu_2\left[r_{12}+\frac{2\nu_1\nu_2}{(100)^2}\right]\right\}$$

Discarding the values that are divided by $(100)^2$, if ν_1 and ν_2 are small,

$$(9)\quad \sigma^2_{(\nu_1-\nu_2)}=1/2N\cdot\{\nu_1^2+\nu_2^2-2r_{12}^2\nu_1\nu_2\}$$

NOTES AND NEWS

A recent report of the Superintendent of Public Instruction of Virginia, Sidney B. Hall, emphasizes the need for training men and women for skilled work which Virginia industries are now forced to have done outside the state. Many manufacturing plants, particularly of silk and rayon, are producing only semi-finished products and are sending these products out of the state to be finished by highly skilled and hence highly paid workers in other states. "Virginia workmen could and should be taught in the interest of the state to perform these highly skilled operations," Mr. Hall says. Approximately 12,000 workers in Virginia received instruction in trade and industrial occupations during the past year. However, this is hardly a beginning when considered in terms of the need for adult industrial education in the state.

"A Background Study of Negro College Students" is the title of a Federal Office of Education bulletin containing a national survey of the social, economic, cultural, academic, and intellectual background factors of nearly 2,000 college students in 33 negro colleges. Significant differences in the scholastic aptitude of students in the different colleges are pointed out. Negro students from schools of the northern and border states make a decidedly higher psychological score than those from the South; also it is revealed that students whose parents belong to certain occupational groups consistently rank high or low in the various measures applied according to the level of the eight occupations studied. The study shows a serious lack of selectivity in negro colleges; a lack of adaptation of the high school program to the varying needs of students; and that negro schools and colleges are not properly providing for the varied vocational needs of negro students, nor are they providing definite educational and vocational guidance programs.

Dr. Harold C. Taylor, formerly Assistant in Psychology, Institute of Human Relations, Yale University, has joined the staff

of the Industrial Relations Division of The Procter & Gamble Company as technical assistant on research.

A study by E. R. Enlow, *A Statistical Slide Rule,* which has been under way for two years has resulted in the production of a genuine calculating slide rule adapted to the statistical calculations in education and psychology. This slide rule is comparable to the engineer's log duplex in versatility and number of scales. Ten of the seventeen scales are standard slide-rule scales which take care of the usual processes of multiplication, division, powers, roots, reciprocals, and logarithms. The remaining seven scales make the statistical slide rule virtually "self-contained" for ordinary computations. They enable the user to read either $1-r^2$ or $\sqrt{1-r^2}$ for a given setting of r, to read areas under the normal curve for given abscissae—or *vice versa,* and to calculate intelligence or other quotients based on two ages given in years and months. An abstract of this study appears in the July, 1934, issue of the *Peabody Journal of Education.*

At the ninth annual meeting of the Midwestern Psychological Association, held at Purdue University, Lafayette, Indiana, May 10-12, 1934, the members elected Dr. John A. McGeoch, University of Missouri, as their next President. Dr. McGeoch had served the previous year as the Secretary-Treasurer of the Association.

In an article titled "This Flying Game" in a recent issue of *School Life,* R. W. Hambrook of the Federal Office of Education points out that "occupations available in connection with the operation of airplanes are just as diverse as those to be found in many other types of industry." Probably no other field of work is so popular with high school boys as that of aviation and it was for the purpose of answering some of the many questions most frequently asked that the study was made. Aviation industry occupations are classified in five groups as follows: Flying Operations, Ground Operations, Manufacturing, Business Phase, and Positions for Women. Women frequently

are employed in aviation as pilots; flying instructors; airport operators, managers, hostesses; and air travel advisers. Preference so far has been given girls who can pilot airplanes and who have been trained as nurses chiefly for the reason that a nurse has an understanding of human nature which enables her to deal with individuals tactfully, under flight conditions.

BOOK REVIEWS

Goodenough, Florence L. *Developmental Psychology.* New York, D. Appleton-Century Co. 1934, xix-619 pp.

In spite of economic depression the last five years have produced an endless succession of textbooks in elementary psychology. Seventy-five per cent of these may be grouped into one class and characterized as compilations of facts and opinions supposedly of psychological interest and reduced to words of one syllable to meet the needs of college students—and many of their professors—who would be lost if they should be required to exhibit a modicum of ratiocination. The remaining twenty-five per cent stand out by reason of lack of conformity to the traditions of textbook writing.

In the present reviewer's opinion Dr. Goodenough's *Developmental Psychology* must be added to the smaller group albeit as a minor member. The book does not present an original point of view as does Kantor's *Survey of the Science of Psychology,* nor is it an elaborate presentation of an older point of view as is Dashiell's *Fundamentals of Objective Psychology.* But the material is organized on an unusual yet logical ontogenetic plan. Here are the traditional subjects of sensory processes, perception, emotion, intelligence, learning, and so on. They are not, however, treated as the static components of the adult human behavioral armamentarium. They are discussed in relation to infancy, childhood and the other conventional stages of growth.

The essential point of view supporting the organization of material is expressed in this quotation: "Psychology is largely concerned with the study of the processes by which early potentialities interact with later experience to form new patterns of behavior, new ways of doing things." Three chapters are devoted to human inheritance and the biological aspects of human growth. Then follow discussions of behavior acquisitions in infancy, early and late childhood, adolescence, adulthood, and old age, with somewhat digressive chapters on intelligence, special defects and abilities, and mental disease. Again may we em-

phasize that the content is that found in the usual elementary texts plus something from the child psychology texts; it is the organization that sets the book apart from most of the others.

It is not as a textbook in elementary psychology that this work will find its only usefulness. Because of much of its content, and because of its organization, it will serve as one of the better texts in child psychology. More particularly is this true in those courses in child psychology which do not require elementary psychology as a prerequisite. Teachers of courses concerned with child behavior in departments of home economics or education would do well to consider carefully the values of this particular book for their needs.

C. M. Louttit
Indiana University

Garrison, Karl C. *The Psychology of Adolescence.* Prentice-Hall, Inc., New York, 1934, xxi-377 pp.

The author has gathered together a considerable amount of data bearing upon the psychology of adolescence and has presented it in its appropriate connections. He has not, however, made his book anything like a compendium of research projects in the field. Indeed he might well have used more of this sort of material without detracting from the value or interest of the book. He has also quoted extensively from the works of psychologists and others who have recorded opinions or made statements relative to the topics discussed in the book. The third and perhaps the most distinguishing characteristic of the book is that the author has, especially in the earlier chapters, briefly outlined a number of principles of general psychology and sought to apply them to the life of the adolescent.

Considering the book from the point of view of its usefulness as a text, the writer's handling of the necessary background of general psychology has been especially fortunate. It is brief enough not to be unduly repetitious to the student who has previously had courses in psychology, but clear enough to be understood by the intelligent reader who has not had such courses.

The selection of the subject matter will appeal to many high school teachers and others interested in the psychological prob-

lems of adolescence. Physical and mental development, motivation, social development, interests, moral and religious growth, personality development, mental disturbances, hygiene, delinquency and guidance are the main subjects treated.

AMOS C. ANDERSON
Ohio University

NEW BOOKS AND PAMPHLETS RECEIVED

Books and pamphlets for review should be sent to James P. Porter, Editor, Journal of Applied Psychology, Ohio University, Athens, Ohio.

An Attempt to Predict the Success of University Freshmen in Their Adjustment to Scholastic Work. Omar C. Held. Edwards Brothers, Ann Arbor, Michigan, 1933. 50 pp.

Child Guidance Clinics. George S. Stevenson and Geddes Smith. The Commonwealth Fund, New York City, 1934. $1.50. 186 pp.

Education and Emergent Man. William C. Bagley. Thomas Nelson & Sons, New York, 1934. 238 pp.

The Effects of Combinations of Handedness and Eyedness on Letter-Position, "Range of Attention," Scores. Irving Anderson and H. R. Crosland. University of Oregon, Eugene, Oregon. 48 pp.

Individual Differences. Frank S. Freeman. Henry Holt & Company, New York, 1934. $2.50. 355 pp.

A Job Analysis of Manufacturing Plants in Minnesota. Charles A. Koepke. Employment Stabilization Research Institute, Vol. II, No. 8. University of Minnesota Press, Minneapolis, Minn. 49 pp.

Manual of Nursing School Practice. Iowa Child Welfare Research Station. University of Iowa, Iowa City, Iowa, 1934.

Mathematics Essential for Elementary Statistics. Helen M. Walker. Henry Holt & Company, New York, 1934. $1.50. 246 pp.

Practice and Variability. Anne Anastasi. Psychological Monograph, Vol. XLV, No. 5. Psychological Review Company, Princeton, N. J., 1934. 55 pp.

Psychology. Third Edition. Robert S. Woodworth. Henry Holt & Company, New York, 1934. $2.50. 546 pp.

Psychology in Education. John Madison Fletcher. Doubleday, Doran & Company, Garden City, N. Y., 1934. $2.50. 524 pp.

The Science of Work. Morris S. Viteles. W. W. Norton & Company, New York, 1934. $4.00. 442 pp.

The Selection of Colour Workers. W. O'D. Pierce. Pitman Publishing Corporation, New York, 1934. $2.00. 134 pp.

Studies in College Examinations. M. E. Haggerty, Chairman, Committee on Educational Research. University of Minnesota, Minneapolis, Minn., 1934. 204 pp.

Time-Limit vs Work-Limit in Learning. Evelyn W. Betts. Mental Measurement Monographs, No. 10, June 1934. Williams & Wilkins Company, Baltimore, Md. 58 pp.

Directory of American Psychological Periodicals

(Continued)

Comparative Psychology Monographs—Baltimore: Johns Hopkins University Press. Subscription $5. 400 pages per volume. Knight Dunlap, Managing Editor. Published without fixed dates, each number a single research.

Genetic Psychology Monographs—Worcester, Mass.: Clark University Press. Subscription $7.00 per volume. Two volumes per year, 600 pages each. Edited by Carl Murchison. Monthly. Each number one complete research. Child behavior, animal behavior and comparative psychology. Founded 1925.

Psychological Abstracts—Eno Hall, Princeton, N. J. American Psychological Association. Subscription $6.00. 700 pages ann. Edited by W. S. Hunter. Monthly. Abstracts of psychological literature. Founded 1927.

Pedagogical Seminary and Journal of Genetic Psychology—Worcester, Mass.: Clark University Press. Subscription $14.00 per annum; $7.00 per volume. 1,000 pages annually (2 volumes). Edited by Carl Murchison and an international cooperating board. Quarterly. Child behavior, animal behavior, and comparative psychology. Founded 1891.

Journal of Social Psychology—Worcester, Mass.: Clark University Press. Subscription $7.00. 500-600 pages annually. Edited by John Dewey and Carl Murchison. Quarterly. Political, racial, and differential psychology. Founded 1929.

Personnel Journal—Baltimore: The Williams & Wilkins Co. Subscription $5.00. 500 pages, approximately. Edited by Walter V. Bingham. Bi-Monthly.

Mental Measurement Monographs—Baltimore: The Williams & Wilkins Company. Subscription $5. 400 pages annually. Edited by Knight Dunlap and Buford Johnson. Published without fixed dates, each number a single research. Founded 1925.

Studies in Psychology and Psychiatry—Baltimore: The Williams & Wilkins Company. Subscription $5.00. 400 pages annually. Edited by Edward A. Pace. Published without fixed dates, each number a single research. Founded 1926.

Journal of Philosophy—New York: 515 West 116th Street. Subscription $4.00. 728 pages per volume. Founded 1904. Bi-weekly. Edited by F. J. E. Woodbridge, Wendell T. Bush, and H. W. Schneider.

Child Development—Baltimore, Md.: Williams & Wilkins Company. Subscription $5.00. 400 pages annually. Edited by Buford Johnson. Quarterly. Founded 1930.

Character and Personality—Durham, N. C.: Duke University Press. Subscription $2.00 per annum. Quarterly. Edited by Robert Saudek (London) and Ernest Seeman (Durham, N. C.). Founded 1932.

Directory of American Psychological Periodicals

American Journal of Psychology—Ithaca, N. Y.: Cornell University. Subscription $6.50. 624 pages annually. Edited by M. F. Washburn, K. M. Dallenbach, M. Bentley, E. G. Boring. Quarterly. General and experimental psychology. Founded 1887.

Journal of General Psychology—Worcester, Mass.: Clark University Press. Subscription $14.00 per annum; $7.00 per volume. 1,000 pages annually (2 volumes). Edited by Carl Murchison and an international cooperating board. Quarterly. Experimental, theoretical, clinical, and historical psychology. Founded 1927.

Psychological Review—Princeton, N. J.: Psychological Review Company. Subscription $5.50. 540 pages annually. Bi-monthly. General Psychology. Edited by Howard C. Warren. Founded 1894.

Psychological Monograph—Princeton, N. J.: Psychological Review Company. Subscription $6.00 per vol. 500 pp. Ed. by Raymond Dodge. Published without fixed dates, each issue one or more researches. Founded 1895.

Psychological Index—Princeton, N. J.: Psychological Review Company. Subscription $4.00. 300-400 pp. Edited by W. S. Hunter. An annual bibliography of psychological literature. Founded 1895.

Psychological Bulletin—Princeton, N. J.: Psychological Review Company. Subscription $6.00. 720 pages annually. Psychological literature. Monthly. Edited by Edward S. Robinson. Founded 1904.

Archives of Psychology—New York: Columbia University. Subscription $6. 500 pp. ann. Ed. by R. S. Woodworth. Published without fixed dates, each number a single experimental study. Founded 1906.

Journal of Abnormal and Social Psychology—Eno Hall, Princeton, N. J. American Psychological Association. Subscription $5. Foreign, $5.25. Editor, Henry T. Moore. Quarterly. 448 pages annually. Abnormal and Social. Founded 1906.

Psychological Clinic—Philadelphia: Psychological Clinic Press. Subscription $3.00. 288 pages. Ed. by Lightner Witmer. Without fixed dates (9 numbers). Orthogenics, psychology, hygiene. Founded 1907.

Psychoanalytic Review—Washington, D. C.: 3617 10th St., N. W. Subscription $6. 500 pages annually. Psychoanalysis. Quarterly. Ed. by W. A. White and S. E. Jelliffe. Founded 1913.

Journal of Experimental Psychology—Princeton, N. J. Psychological Review Company. 500 pages annually. Experimental. Subscription $6.00. Bi-monthly. Edited by S. W. Fernberger. Founded 1916.

Journal of Applied Psychology—Crawfordsville, Ind.: R. R. Donnelley & Sons Co. Subscription $5.00. 600 pages annually. Bi-monthly. Edited by James P. Porter, Ohio University, Athens, Ohio. Founded 1917.

Journal of Comparative Psychology—Baltimore: The Williams & Wilkins Co. Subscription $5.00 per volume of 450 pages. Three volumes every two years. Edited by Knight Dunlap and Robert M. Yerkes. Founded 1921.

(Continued on Third Cover)